Audit Plan and Evaluation Framework for

Pennsylvania Act 129

Energy Efficiency and Conservation Programs

Prepared by The Statewide Evaluation Team:

GDS Associates, Inc., Nexant, & Mondre Energy

Contracted Under the Pennsylvania Public Utility Commission's

RFP 2009-1 for the Statewide Evaluator

Revised: November 4, 2011

Table of Contents

List of Tables

List of Figures

List of Acronyms

B/C Ratio: Benefit-Cost Ratio

BTUS: Bureau of Technical Utility Services

C&I: Commercial and Industrial

CEEP: Conservation, Economics, and Energy Planning

CFL: Compact Fluorescent Light

CMP: Custom Measure Protocol

Cv: Coefficient of Variance

CSP: Conservation Service Provider

The Commission: The Pennsylvania Public Utility Commission

DLC: Direct Load Control

DR: Demand Response

DSM: Demand Side Management

EDC: Electric Distribution Company

EE: Energy Efficiency

EE&C Plan: Energy Efficiency and Conservation Plan

EER: Energy Efficiency Resource

EM&V: Evaluation, Measurement and Verification

HVAC: Heating, Ventilating, and Air Conditioning

IMP: Interim Measure Protocol

IPMVP: International Performance Measurement and Verification Protocol

kW: Kilowatt

kWh: Kilowatt-Hour

M&V: Measurement and Verification

NPV: Net Present Value

NTG: Net-to-Gross, Net-to-Gross Ratio

PA PUC: Pennsylvania Public Utility Commission

PUC: see PA PUC

PY: Program Year

SEM: Simple Engineering Model

SSMVP: Site Specific M&V Plan

SWE: Statewide Evaluator

SWE Team: Statewide Evaluation Team

TOU: Time of Use

TRC: Total Resource Cost

TRM: Technical Reference Manual

TWG: Technical Working Group

VFD: Variable Frequency Drive

Please see Appendix A for definition of terms.

Executive Summary

This updated Audit Plan has been prepared by the Statewide Evaluation (SWE) Team retained by the Pennsylvania Public Utility Commission (PUC or the Commission) to audit the evaluation activities and program savings of the seven Pennsylvania Electric Distribution Company (EDC) Energy Efficiency and Conservation (EE&C) Plans, approved by the Commission, in accordance with the directive of Act 129. The Audit Plan contains detailed expectations for EDC evaluation, measurement and verification (EM&V) protocols and plans as well as the potential audit activities to be conducted by the SWE Team. Specifically, the Audit Plan addresses the following:

- Specific audit activities to be undertaken by the SWE Team;
- Review of EDC EE&C program plans and impact evaluation expectations;
- Plans for developing and implementing annual updates to the Technical Reference Manual (TRM);
- The role of the Technical Working Group (TWG) to update and clarify the TRM;
- EDC gross energy and demand impact evaluations with EM&V plan guidelines based on specified savings and sampling protocols;
- EDC process evaluations with guidelines for creating and conducting surveys;
- Review of EDC cost-effectiveness evaluations with guidelines for following the Total Resource Cost (TRC) Test described in the Commission's TRC Orders of June 2009 and August 2011;
- Data tracking and reporting guidelines; and
- Deadlines for evaluation activities, audit activities, and reporting.

As noted in the contract between the SWE and the PA PUC, this Audit Plan is considered to be a living document that can be revised on a regular basis throughout the contract term.[1] A list of changes made at each update can be found in the Executive Summary. The Statewide Evaluation Team, with the approval of the Bureau of Technical Utility Services (BTUS)[2], can update the Audit Plan as needed to adjust for changes in EDC plans. It is anticipated that such updates will occur at such times that will allow EDCs ample time to adjust their data collection and reporting methods and meet their annual reporting requirements.

The Statewide Evaluation team is committed to working collaboratively with the PA PUC staff, the seven EDCs subject to Act 129, and the EDC evaluation teams to develop and implement the most efficient and effective evaluation, measurement and verification processes for the Act 129 energy efficiency and demand response programs. To date, this collaborative process for the audit activities conducted by the SWE has allowed for comprehensive review and assessment of program savings within the constraints of the available impact and process evaluation budget for the Act 129 programs.

[1] See the Act 129 Statewide Evaluator RFP 2009-1, page 33, Section IV-3(e).
[2] Formerly known as the Conservation, Economics, and Energy Planning (CEEP) Bureau.

November 2011 Updates to the Audit Plan

A revision to the SWE Audit Plan was issued on July 2011. Subsequently, the SWE received several comments from the EDCs, prompting a second revision in 2011. Major areas or improvement since the July 2011 Audit Plan include:

- Changed all references to CEEP to BTUS.
- Changed all references to Allegheny Power to West Penn Power.
- Added Section 1.4 regarding guidance memos.
- Updated TRC-related issues according to the 2011 TRC Order, including clarifying the use of NTG ratios for planning purposes and modifying the definition of free-ridership and spillover.
- Corrected due dates of SWE quarterly and annual reports.
- Added a clarification that impact evaluation results, as they are made available, will be used to update the TRM during the annual TRM update process.

July 2011 Updates to the Audit Plan

The original SWE Audit Plan was issued on December 1, 2009. This version of the SWE Audit Plan incorporates updates stemming from issues discussed with BTUS and the EDCs and modifications to the audit process in an effort to streamline the SWE's activities. Major areas of improvement since the 2009 Audit Plan include:

- Moved "Table of Contents", "List of Tables", "List of Figures" and "List of Acronyms" to the front immediately following the title page. Moved "Roles and Responsibilities" to Section 1. Moved "Baseline Studies" to Section 3. Moved "Guidance on Impact Evaluation", "Guidance on Process Evaluation", "Guidance on Cost-Effectiveness Analysis" and "Sampling and Uncertainty Protocol" to Section 4. Moved "Audit Objectives and Process" to Section 5. Moved "Data Tracking and Reporting Guidelines" to Section 6.
- Created "Research Objectives", "Guidance on Calculating Claimed Savings", "Calculating Verified Gross Savings", "Calculating Verified Net Savings", "Interim Measure Protocols", "Custom Measure Protocols", "Reporting Savings", "Dynamic Sampling Methodology", "Pennsylvania Act 129 SharePoint Site", and "Public Accessible Website Data Requirements" sections.
- Updated program summaries based on most recent EE&C plans filed with the Commission.
- Clarified Section 3 and 4 to distinguish between responsibilities of the program administrator (EDC, EDC Conservation Service Providers, and/or EDC Implementer) and the EDC evaluator.
- Provided framework for interim protocol approval process.
- Provided framework for custom measure protocols.
- Updated sampling and uncertainty methodology.
- Updated expected deadlines for key deliverables.

Clarifying the TRM

The first version of the TRM, approved in May 2009, categorizes all measures and/or programs into three categories: deemed measures, partially deemed measures, and custom measures. Deemed

measures require no additional measurement to determine savings; only, verification of installation must be conducted in order to effectively estimate the realization rate of these measures. Partially deemed measures include both stipulated and open variables, thereby requiring the verification of installation and the measurement of certain variables in order to define and evaluate the energy and peak demand savings. Lastly, custom measures are outside the scope of the TRM altogether. Through the collaborative efforts of the TWG, led by the SWE Team, the TRM will be updated annually. These annual updates will involve the Commission's typical stakeholder process which follows the Tentative Order/Final Order Commission procedure. Program year one TRM updates, completed during 2009 and 2010, included pertinent issues such as defining variables that required clarification for consistent interpretation, and clarifying the appropriate application of the TRM stipulated values. Future updates of the TRM will involve prospective measure clarifications and the addition of new measures. This is discussed in Section 3.2.1 of the Audit Plan.

Approval procedures for measures not in the TRM

According to the 2009 TRM Order[3], any measure not included in the current version of the TRM falls into one of two categories; "potential standard measures not currently included in the TRM" that can be incorporated into future versions of the TRM, and "custom measures" that are outside the scope of the TRM. The first category of measures is addressed by the interim measure protocol process and the second category by the custom measure protocol process. It is the intention of the SWE to continue to work with the EDCs and EDC evaluators to develop interim measure protocols for new measures that the EDCs would like to offer in their programs. The SWE will maintain a catalog of interim measure protocols on the SWE SharePoint site in order to maintain a database for new measures that can be offered in EDC programs. More information on this process is included in Section 3.2.2 and Section 3.2.3 of this Audit Plan.

Establishing the value of the programs through impact evaluation

The emphasis of this Audit Plan is on measuring annual program energy and demand impacts, given that these impacts are critical to the success of each EDC in attaining the energy and demand savings targets outlined in Act 129. The value of the EDC programs is primarily established through the impact evaluations that verify the resource acquisition savings. Evaluation activities identify, document, quantify, and monetize these impacts. The EDC evaluators are expected to conduct program-specific data collection and reporting, as outlined in their respective EM&V plans, that supports unbiased independent estimations of verified gross[4] energy and demand impacts for all programs.

These energy and demand impacts are to be reported in annual reports written by the EDCs and submitted to the Commission by November 15th of each year[5]; the timetable for deliverables is discussed in Section 1.2. All of the issues pertaining to the impact evaluation process are detailed in Section 4.1 of the Audit Plan. The audit activities of the SWE will be implemented throughout the

[3] Order Entered on June 1, 2009, at Docket No. M-00051865. Page 9-10.
[4] Note: Pending Commission decision, EDCs may be required to evaluate the net impacts of each program in future program years.
[5] Secretarial Letter issued on May 25, 2011, at Docket No. M-2008-2069887. Page 1-2.

evaluation process. For more information on what is to be audited, how it will be audited, and the possible audit outcomes, please refer to Section Audit Objectives and Process of the Audit Plan as well as the Audit Activities Checklist provided in Appendix D.

Establishing the effectiveness of the programs' implementation through process evaluation

The objective of the process evaluation activities is to identify best-practice implementation strategies in order to continually improve the efficiency and effectiveness of EDC program delivery. The process evaluations will be conducted by the EDC evaluators on an on-going, as-needed basis. This Audit Plan contains the SWE guidelines for these process evaluations. These expectations include:

- Determining adequate timing of process evaluation activities
- Defining program design
- Identifying market barriers and implementing effective strategies for overcoming them
- Assessing the efficiency and effectiveness of program administration
- Reviewing program implementation and delivery and assessing its effectiveness from the perspective of all program stakeholders
- Understanding the market response to EE&C programs implemented

The process evaluations will be reviewed by the SWE in order to identify best-practices and areas for improvement across all seven EDCs. The findings will be shared in biannual workshops conducted by the SWE Team so that best-practices can be implemented immediately. The details of the process evaluation activities, as specified by the SWE, are discussed in Section 4.2 of this Audit Plan.

Establishing the cost-effectiveness of the programs

Given that budgets are constrained and the savings targets are aggressive, the cost-effectiveness of the EDC programs is paramount. As detailed in the Commission's TRC Order approved in June 2009[6], the EDC programs must pass the TRC benefit-cost ratio at the portfolio (i.e., EDC specific EE&C Plan) level to be considered cost-effective. This Audit Plan outlines the directives and stipulations of the TRC calculations as defined by the Commission in the 2009 and 2011 TRC Orders[7]. The Audit Plan also clarifies some of the assumptions so that the TRC calculations will be implemented consistently across all EDCs. And, although the programs must past the TRC test at the portfolio level, the SWE will review all costs and calculations at the individual program level in order to draw conclusions regarding the cost-effectiveness of programs in terms of the trade-offs between overall budgets and the attainment of Act 129 goals. The details of the cost-effectiveness evaluation, as specified by the SWE, are discussed in Section 4.3 of the Audit Plan.

[6] *Implementation of Act 129 of 2008 – Total Resource Cost Test (TRC) Test*. Pennsylvania Public Utility Commission Docket NO. M-2009-2108601. June 18, 2009.
[7] *Implementation of Act 129 of 2008 – Total Resource Cost Test (TRC) Test.* Pennsylvania Public Utility Commission Docket No. M-2009-2108601. August 2, 2011.

Auditing of EDC evaluations

The SWE Team has been retained by the PA PUC to conduct a statewide audit and assessment of the seven EDC EE&C program plans. The SWE audit activities include:

- Standardizing impact, process, and cost-effectiveness evaluation protocols;
- Reviewing EM&V evaluation plans for completeness and consistency;
- Auditing EM&V activities for quality control, accuracy, and uncertainty mitigation; and
- Assessing the achievements of each EDC and the EDCs combined in order to evaluate:
    - Progress towards accomplishing Act 129 goals;
    - Cost-effectiveness of programs; and
    - Effectiveness of program implementation and delivery.

The SWE will submit annual audit reports which will provide the findings, insight, and feedback required to improve or sustain the EDC EE&C programs in order to meet the directives of Act 129 and to verify or modify reported impact evaluations filed by the EDCs. The SWE must ensure that the program impacts reported by the EDCs are valid and represent the true EE&C program effects. The SWE reserves the right to conduct audit and evaluation activities at a level which guarantees that the evaluation findings are both accurate, and have the specified level of confidences and precision established in this Audit Plan. The Audit Process is discussed in Section 5 of this Audit Plan.

Data requirements, deliverables and timetables

In order for the SWE Team to carry out its responsibilities, it is necessary that certain data be provided to the SWE. Additionally, it is imperative that all deliverables be distributed according to the pre-specified timetable in order for all parties involved to carry out their individual responsibilities in accordance with Act 129 and the legislative timetable. The specific data requirements that are expected to be provided by the EDCs or EDC evaluators are laid out in Section 6 of this Audit Plan. All deliverables and dates are specified in Section 6.3.

Goals and Objectives

The primary objective for each EDC is to reach the level of savings specified in Act 129 in a meaningful, efficient, and cost-effective manner. It is the desire of the SWE to continue to work closely with the PUC and EDCs in order to develop and implement an evaluation and audit process that will produce significant and standardized impact results at the lowest reasonable cost, so that more funds can be allocated to customer-centric savings activities. However, as the SWE, it is our duty to verify that the evaluations are accurate and represent the actual impacts of the EE&C programs with a certain level of precision and confidence deemed appropriate by the SWE Team and PUC.

1 Introduction and Purpose of the Audit Plan

This Audit Plan has been developed by the SWE Team pursuant to the evaluation requirements under Act 129 and the EE&C Program Implementation Order. Included in the Audit Plan are guidelines and expectations for the seven Pennsylvania EDCs whose program plans were approved by the PA PUC to promote the goals and objectives of Act 129. The EDCs are Duquesne Light Company, Metropolitan Edison Company, PECO Energy Company, Pennsylvania Electric Company, Pennsylvania Power Company, PPL Electric Utilities Corporation, and West Penn Power Company. The SWE Team has been retained by the PA PUC to fulfill the statewide evaluation requirements of Act 129 and the EE&C Program Implementation Order. In accordance with the RFP for the Statewide Evaluator contract awarded by the PA PUC in September of 2009, the SWE Team will:

- Develop the evaluation framework specifying:
  - Expectations for evaluation activities
  - Savings protocols
  - Standard data to be collected
  - Databases for the evaluation of program benefits and results to be used across all EDC service territories
- Perform ongoing audits of the reported impact and cost-effectiveness of each EDC EE&C Plan
- Perform a five year independent market assessment that includes:
  - Review and approval of statewide market assessments, characterizations, and TRM updates
  - Statewide Market Potential Study for additional energy and load reductions

This updated Audit Plan serves as an evaluation framework that outlines the expected metrics, methodologies and guidelines for measuring performance by detailing the processes that should be used to evaluate the programs sponsored by the EDCs throughout the state of Pennsylvania. It also sets the stage for discussions amongst a Technical Working Group (TWG) of the EDCs, their evaluators, the SWE Team and the PUC. During these discussions, the TWG will clarify and interpret the TRM, recommend additional prescriptive measures to be included in the TRM, and define guidelines for acceptable measurement protocols for custom measures in order to mitigate risks to the EDCs. Our common goal requires kWh and kW savings to be clearly defined, to be auditable, and to provide sound engineering bases for estimating verified gross energy savings.

This Audit Plan addresses the following:

- Savings protocols
- Metrics and data formats
- Guidance on claiming savings
- Guidance on impact evaluation procedures
- Guidance on process evaluation procedures

- Guidance on cost-effectiveness analysis
- Required reporting formats
- Data management and quality control guidelines
- Data tracking and reporting systems
- The SWE Team's public website
- Cost-effectiveness calculation requirements
- Description of how the SWE Team will audit the evaluations performed by EDCs
- Criteria the SWE Team will use to review and assess EDC evaluations

This plan also outlines the schedule of activities to be conducted by the SWE Team in order to audit each EDC's evaluations and to assess the progress made by the EDCs, both individually and collectively, towards the attainment of the energy and demand savings targets set by Act 129.

To fulfill these requirements, the SWE Team will work in collaboration with the seven Pennsylvania EDCs, their respective evaluation teams, and the PA PUC staff, in order to develop meaningful and uniform procedures so that the performance of each EDC's EE&C program is verifiable, reliable and meets the objectives of the Act under which it was developed.

The contents of this Audit Plan are to be adopted immediately. Any updates made to the Audit Plan are intended to clarify and memorialize decisions made through other means (e.g. Orders, Secretarial Letters, Guidance Memos, etc.).

1.1 Act 129 Requirements for the Statewide Evaluation

As a part of Act 129 and the Implementation Order, the PA PUC sought a Statewide Evaluator to evaluate the EDCs' energy efficiency and conservation plans by issuing an RFP on April 21, 2009. The Commission's RFP specified the services to be provided by the SWE in accordance with the requirements of Act 129 and the Implementation Order. These services included monitoring and verifying EDC data collection, quality assurance processes, and performance measures by customer class. The SWE is also responsible for auditing the results of each EDC plan on an annual basis and updating the entire energy efficiency and conservation program goals as a whole in 2013. The annual audits will include an analysis of plan and program impacts (demand and energy savings) and cost-effectiveness. The SWE is to report results and provide recommendations for plan and program improvements. The RFP states that the SWE will produce an accurate assessment of the future potential for energy savings through market potential assessments. The RFP also specifies that these programs are being implemented pursuant to Act 129 of 2008 and that the evaluations shall be conducted within the context of the Implementation Order and Act 129.[8]

[8] The Commission has been charged by the Pennsylvania General Assembly pursuant to Act 129 of 2008 ("Act 129") with establishing an energy efficiency and conservation program. 66 Pa.C.S. §§ 2806.1 and 2806.2. The energy efficiency and conservation program requires each EDC with at least 100,000 customers to adopt a plan to reduce energy demand and consumption within its service territory. 66 Pa.C.S. § 2806.1. In order to fulfill this obligation, on January 16, 2009, the Commission entered an Implementation Order at Docket No. M-2008-

The PA PUC selected GDS Associates, Inc. as the prime contractor, with Nexant and Mondre Energy as subcontractors. The PA PUC authorized the SWE Team to begin work on September 3, 2009. The scope of work approach proposed by the SWE Team met all of the requirements for tasks and deliverables that are listed in the Commission's RFP. The scope of work for the statewide evaluation was developed to provide the required level of verification described in the RFP and at the pre-bid meeting. The approach will audit the verification work conducted by the EDC evaluators.

The SWE audit will include an analysis of each EDC plan from a process, impact, and cost-effectiveness standpoint. The SWE Team will provide quarterly process updates as well as biannual best practice workshops with the EDCs. The annual reports produced will provide the Commission with recommendations for EE&C Plan and program improvements. Additionally, the SWE Team will provide a public web-accessible database and reporting system for the Commission's website so that the general public may be kept abreast of the impacts of the EE&C Plans by program and sector. As the SWE, we will also produce an accurate assessment of the future potential for energy savings through a market assessment study. While all of these tasks are related, they each have distinct goals:

- Impact evaluation reviews seek to *quantify t*he energy, demand, and possible non-energy impacts[9] that have resulted from demand-side management (DSM) program operations;
- Process evaluations seek to *describe* how well those programs operate and to *characterize* their efficiency and effectiveness; and
- Market Characterization and Assessment seeks to *determine* the attitude and awareness of market actors, measure market indicators, and identify barriers to market penetration[10].

With this Audit Plan, the SWE has outlined expectations and guidelines for EDC EM&V plans and activities, and provided a description of the audit activities that will be carried out by the SWE Team in order to audit the EE&C program impacts reported by the EDCs and verified by EDC evaluators, and to assess the progress towards the completion of the Act 129 energy conservation goals.

1.2 Roles and Responsibilities

According to the RFP issued for the Statewide Evaluation Contractor, the following tables delineate the roles and responsibilities for the EDCs, the SWE Team, and the PUC, by tasks and deliverable, falling under the following categories:

2069887. As part of the Implementation Order and Act 129, the Commission issued an RFP for a Statewide Evaluator (on April 21, 2009) to evaluate the EDCs' energy efficiency and conservation programs.

[9] Note: No non-energy benefits have been specified for the Act 129 programs as of program year one.

[10] Note: The SWE is responsible for conducting a Statewide Market Assessment in 2012. However, the EDCs are responsible for any Baseline Studies, NTG Studies, or Market Characterization Studies required prior to 2012.

- Statewide studies
- Planning and management
- Process evaluation
- Databases
- Primary data collection and impact analyses
- Independent data collection and impact analyses
- EDC plan review
- Cost-effectiveness analyses
- Reporting
- Best practices
- Other

Table 1-1: Roles and Responsibilities - Statewide Studies

Task and/or Deliverable	EDC	SWE	PUC
Coordinated Statewide Market Assessments, Characterizations (e.g. Baseline studies).	XX	XX[11]	
Review and approval of Statewide Market Assessments, Characterizations and TRM updates.		XX	XX
Statewide Market Potential Study for additional energy and load reductions after May 31, 2013.		XX	
Annual updates to TRM.		XX	
Approval of TRM updates.			XX
Develop and conduct NTG studies.	XX		
Coordinate the development of and approve the methodologies for EDC NTG studies.		XX	

Table 1-2: Roles and Responsibilities - Planning and Management

Task and/or Deliverable	EDC	SWE	PUC
EDC impact and process evaluation plans; including database and reporting protocols, survey templates, and schedules[12].	XX		
Review and approval of EDC evaluation plans.		XX	XX
Develop Statewide audit evaluation plan, including cost-effectiveness plan, verification approach (with sampling plan), and schedules.		XX	
Review and approval of statewide audit evaluation plan.			XX
Coordination of all EDC evaluation efforts.	XX	XX	
Coordination of statewide impact and cost-effectiveness evaluation efforts.		XX	

[11] Note: The baseline study was not originally part of the SWE's responsibilities listed in the RFP and was added at a later date.

[12] Note: While the SWE will require specific data reports and reporting guidelines for transferring information to the SWE database and tracking system, the EDCs are free to develop their own internal database and reporting protocols.

Table 1-3: Roles and Responsibilities - Process Evaluation

Task and/or Deliverable	EDC	SWE	PUC
Program Process Evaluations.	XX		
Customer and trade-ally satisfaction surveys and reports[13].		XX	

Table 1-4: Roles and Responsibilities - Databases

Task and/or Deliverable	EDC	SWE	PUC
Design, implementation and maintenance of EDC primary program tracking database(s) with project and program data.	XX		
Establishing and implementing quality control reviews of EDC program tracking databases.		XX	
Statewide Data Management and Quality Control. Design, implementation and maintenance of statewide database of program, portfolio, EDC and statewide energy and demand savings and cost-effectiveness reporting; provide public web accessible database and reporting system.		XX	
Approval of statewide database and reporting plan.			XX

Table 1-5: Roles and Responsibilities - Primary Data Collection and Impact Analyses

Task and/or Deliverable	EDC	SWE	PUC
First year primary data collection and site baseline and ex-post verification for EE projects.	XX		
Persistence of savings analysis for Custom Measures[14]: primary data collection.	XX		
Analyses and documentation of project, program and portfolio gross and net energy and demand savings.	XX		

Table 1-6: Roles and Responsibilities - Independent Data Collection and Impact Analyses

Task and/or Deliverable	EDC	SWE	PUC
Quality control and due diligence. Inspections of project sites and review of primary data and analyses, preparation of verified achieved versus claimed savings		XX	

[13] Note: Typically, these types of surveys are included as part of a Process Evaluation; therefore, the EDCs should expect to survey all stakeholders for their evaluations. However, additional surveys conducted by the SWE may be required for the process evaluation audit. The SWE will coordinate with the EDCs to minimize logistical issues.

[14] Note: Persistence is always an issue. However, most of the prescriptive measures are deemed which implies two things: (1) the savings are pre-determined and (2) the equipment is typically standard with less variability in usage and persistence (or longevity of performance) than custom measures. Therefore, in the interest of upholding the TRM and managing EM&V costs, the "persistence issue" will primarily be limited to custom measures.

Table 1-7: Roles and Responsibilities - EDC Plan Review

Task and/or Deliverable	EDC	SWE	PUC
Review of filed EDC plans and provide advice to Commission staff on ability of plans to cost effectively meet targets.		XX	
Review of EDCs' EM&V plans and provide advice to Commission staff on ability of plans to adequately measure energy savings.		XX	

Table 1-8: Roles and Responsibilities - Cost-Effectiveness Analyses

Task and/or Deliverable	EDC	SWE	PUC
EDC and participant cost reporting and EDC cost-effectiveness analyses.	XX		
Review, audit and analysis of EDC and participant cost reporting; independent cost-effectiveness analyses, statewide analyses.		XX	

Table 1-9: Roles and Responsibilities - Reporting

Task and/or Deliverable	EDC	SWE	PUC
EDC status reports to statewide evaluator on program implementation.	XX		
Statewide evaluator status reports to Commission on program implementation.		XX	
EDC quarterly and annual reporting of EE program and portfolio net[15] and gross impacts, as applicable, and cost-effectiveness, and EDC progress in reaching targets.	XX		
Statewide quarterly and annual report of program and portfolio results: net and gross impacts and cost-effectiveness and EDC progress in reaching targets.		XX	
Review PUC evaluation Contractor's statewide and EDC quarterly report of EE program and portfolio results.			XX
Review EDC and PUC contractor's Annual Report on EE Programs: net and gross savings impacts, cost-effectiveness, and EDC progress in reaching targets.			XX

Table 1-10: Roles and Responsibilities - Best Practices

Task and/or Deliverable	EDC	SWE	PUC
Participation in at least semi-annual impact evaluation process review and improvement workshops as needed.	XX	XX	XX
Preparation of best practices recommendations for improvements to impact evaluation processes.	XX	XX	
Prepare best practices recommendations for program modifications and improvements.	XX	XX	

[15] Note: Although NTG analysis may be done once a year, the ratios can be used to estimate net impacts on a quarterly basis as gross savings estimates roll-in. Therefore, once NTG ratios are established, then net savings can be estimated just as frequently as gross savings are reported. Net savings, however, will only be calculated and applied if ordered by the Commission.

Table 1-11: Roles and Responsibilities - Other

Task and/or Deliverable	EDC	SWE	PUC
Prepare materials and reports in support of Commission analysis of efficiency programs		XX	
Organize and conduct periodic statewide workshops on evaluation results of EE programs		XX	

1.3 Research Objectives

This Audit Plan seeks to answer research objectives related to impacts and processes for different audiences interested in the results of the Act 129 Program. These objectives are detailed in the following table.

Table 1-12: Audit Plan Research Objectives

Target Audience	Impact Questions	Process Questions
Pennsylvania Legislature	• Did the EDCs meet statutory targets of 1% energy savings by May 31, 2011, and 3% energy savings and 4.5% demand savings by May 31, 2013? • Were the EDC EE&C plans implemented in a cost-effective manner in accordance with the TRC test? • Did the EDC EE&C plans produce real energy and demand savings?	• Which programs were the most successful and why? • Which programs were the most cost-effective and why? • How can EDCs improve programs in order to meet statutory targets?
Pennsylvania PUC	• What level of program energy savings was verified for each EDC and how does this compare to planning estimates and savings reported in quarterly and annual reports? • What assumptions related to energy and demand savings need to be updated for the next round of Act 129? • What were the largest sources of uncertainty identified by EDC evaluators related to energy and demand savings and cost-effectiveness?	• Why did planning estimates and reported gross savings differ from verified gross savings? • Considering differences in planning estimates, reported gross savings, and verified gross savings, how can program planning and reporting be improved? • What sectors have the greatest potential for additional energy and demand savings?

Pennsylvania EDCs	• What factors contributed to differences between planning estimates and reported gross savings? • What factors contributed to differences between reported gross savings and verified gross savings? • Which programs performed the best and which programs require modification or consideration for elimination based on results of the evaluators?	• What changes can the EDCs adopt to minimize differences between planning estimates, reported gross savings, and verified gross savings? • What changes can the EDCs adopt to increase cost-effectiveness of the programs and portfolios?

1.4 Guidance Memos

This Audit Plan was originally developed to provide an overarching framework for Act 129 programs and may not address all nuances uncovered throughout the actual implementation process. For such issues, the SWE has developed guidance memos as a means to memorialize decisions made. All guidance memos will be issued by the SWE and undergo an iterative review process with input from EDC teams and the BTUS staff. All guidance memos will be posted on the SWE Share Point site once approved.

It must be noted that guidance memos do not reflect the opinions, regulations, or rulings of the Commission and, therefore, are not binding on the Commission.

2 EDC Portfolio Overview

The SWE Team reviewed the Act 129 EE&C Plans submitted by the Pennsylvania EDCs. The primary purpose of the review was to determine the programs expected to have the most impact on energy and demand savings. This review provided guidance in determining where to focus the level of resources and effort in M&V and impact evaluations to those programs whose energy savings and demand savings best meet the goals of each EDC and Act 129. The following tables show the breakdown of projected annual kWh and kW savings for the EDCs' EE&C portfolios by sector, as they are estimated in the current versions of the EDC specific EE&C Plans.

2.1 EDC Program Summaries

Table 2-1: West Penn Power – Summary of Portfolio Energy and Demand Savings

West Penn Power *June 30, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	20,938,650	3,496.00	20,938,650	3,496.00	20,938,650	3,496.00	20,938,650	3,496.00
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	12,100	1.44	72,891	9.26	168,304	26.68	264,458	52.56
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	979	0.15	7,903	1.28	12,902	2.03	17,791	2.75
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	8,307	1.91	66,236	14.73	167,626	38.87	222,764	50.82
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	399	0.08	24,023	58.39	59,529	86.85	76,878	102.23
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	2,842	0.26	38,903	3.04	57,519	5.80	63,997	7.96
EE&C Plan Total - Cumulative Projected Savings	**24,627**	**3.84**	**209,956**	**86.70**	**465,880**	**160.23**	**645,888**	**216.33**
Percent Reduction From Baseline	0.1%	0.1%	1.0%	2.5%	2.2%	4.6%	3.1%	6.2%
Commission Identified Goal			209,387				628,160	157.32
Percent Savings Due to Portfolio Above or Below Commission Goal			0.3%				2.8%	1.9%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

Table 2-2: Duquesne Light – Summary of Portfolio Energy and Demand Savings

Duquesne Light *July 1, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	14,085,512	2,518.00	14,085,512	2,518.00	14,085,512	2,518.00	14,085,512	2,518.00
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	19,127	9.17	56,358	33.32	93,822	57.53	131,520	81.80
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	4,294	1.75	12,881	5.25	21,468	8.75	30,055	12.25
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	12,100	2.24	36,413	9.30	60,838	16.36	85,375	23.43
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	37,136	7.25	111,580	25.34	186,915	43.42	261,586	61.51
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	8,973	2.88	26,920	8.65	44,867	14.42	62,814	20.19
EE&C Plan Total - Cumulative Projected Savings	**81,630**	**23.29**	**244,152**	**81.86**	**407,910**	**140.49**	**571,350**	**199.18**
Percent Reduction From Baseline	0.6%	0.9%	1.7%	3.3%	2.9%	5.6%	4.1%	7.9%
Commission Identified Goal			140,885				422,565	113.00
Percent Savings Due to Portfolio Above or Below Commission Goal			73.3%				35.2%	24.3%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

Table 2-3: Met-Ed – Summary of Portfolio Energy and Demand Savings

Met-Ed *September 21, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	14,865,036	2,644.00	14,865,036	2,644.00	14,865,036	2,644.00	14,865,036	2,644.00
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	16,020	7.18	92,760	38.95	169,477	70.17	237,994	82.51
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	1,315	0.12	6,385	0.60	11,392	1.06	16,080	1.50
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	7,859	2.95	45,596	16.21	83,332	29.47	121,051	42.72
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	1,878	0.60	11,768	12.59	21,658	16.59	31,548	9.58
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	3,478	0.87	21,403	5.13	39,300	9.36	51,255	12.69
EE&C Plan Total - Cumulative Projected Savings	**30,550**	**11.71**	**177,912**	**73.47**	**325,159**	**126.65**	**457,928**	**148.99**
Percent Reduction From Baseline	0.2%	0.4%	1.2%	2.8%	2.2%	4.8%	3.1%	5.6%
Commission Identified Goal			148,650				445,951	118.98
Percent Savings Due to Portfolio Above or Below Commission Goal			19.7%				2.7%	6.4%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

Table 2-4: Penelec – Summary of Portfolio Energy and Demand Savings

Penelec *July 1, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	14,399,289	2,395.00	14,399,289	2,395.00	14,399,289	2,395.00	14,399,289	2,395.00
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	14,164	3.48	83,330	29.40	152,472	54.72	216,253	63.39
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	316	0.01	1,255	0.08	2,200	0.14	3,045	0.20
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	8,503	2.86	49,662	16.06	90,821	29.26	131,962	42.45
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	2,483	0.62	14,964	13.73	27,446	16.84	39,928	9.95
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	3,574	0.78	22,026	4.72	40,459	8.64	51,594	11.52
EE&C Plan Total - Cumulative Projected Savings	**29,040**	**7.75**	**171,237**	**63.98**	**313,398**	**109.59**	**442,782**	**127.51**
Percent Reduction From Baseline	0.2%	0.3%	1.2%	2.7%	2.2%	4.6%	3.1%	5.3%
Commission Identified Goal			143,993				431,979	107.78
Percent Savings Due to Portfolio Above or Below Commission Goal			18.9%				2.5%	1.7%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

Table 2-5: Penn Power – Summary of Portfolio Energy and Demand Savings

Penn Power *July 1, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	4,772,937	980.00	4,772,937	980.00	4,772,937	980.00	4,772,937	980.00
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	4,358	1.08	25,490	6.81	46,590	11.87	66,227	14.94
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	93	0.00	357	0.02	624	0.04	859	0.06
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	3,064	1.17	17,861	6.53	32,657	11.89	47,449	17.25
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	749	0.27	4,862	16.61	8,975	17.95	13,088	4.28
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	1,170	0.30	7,277	1.80	13,384	3.30	16,741	4.41
EE&C Plan Total - Cumulative Projected Savings	**9,434**	**2.82**	**55,847**	**31.77**	**102,230**	**45.04**	**144,364**	**40.93**
Percent Reduction From Baseline	0.2%	0.3%	1.2%	3.2%	2.1%	4.6%	3.0%	4.2%
Commission Identified Goal			47,729				143,188	44.10
Percent Savings Due to Portfolio Above or Below Commission Goal			17.0%				0.8%	2.1%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

Table 2-6: PECO – Summary of Portfolio Energy and Demand Savings

PECO *July 1, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	39,248,500	7,889	39,248,500	7,889	39,248,500	7,889	39,248,500	7,889
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	90,796	5.83	239,542	47.50	396,239	86.40	502,833	121.66
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	6,096	0.41	22,239	1.45	49,479	3.23	79,660	5.31
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	14,772	3.31	137,056	104.79	257,529	213.08	385,183	322.82
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	11,800	2.35	80,011	15.82	148,222	29.28	216,792	42.93
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	0	0.00	110,194	11.55	110,516	11.98	111,097	12.73
EE&C Plan Total - Cumulative Projected Savings	**123,464**	**11.89**	**589,042**	**181.11**	**961,985**	**343.97**	**1,295,565**	**505.44**
Percent Reduction From Baseline	0.3%	0.15%	1.5%	2.30%	2.5%	4.36%	3.3%	6.41%
Commission Identified Goal			393,850				1,181,550	355.00
Percent Savings Due to Portfolio Above or Below Commission Goal			49.6%				9.6%	42.4%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.
N/A: Not available.

Table 2-7: PPL – Summary of Portfolio Energy and Demand Savings

PPL *July 1, 2009 EE&C Plan*	Program Year 2009		Program Year 2010		Program Year 2011		Program Year 2012	
	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved	MWh Saved	MW Saved
Baseline[1]	38,214,368	6,591.95	38,214,368	6,591.95	38,214,368	6,591.95	38,214,368	6,591.95
Residential Sector (exclusive of Low-Income) - Cumulative Projected Portfolio Savings[2]	29,647	3.95	153,260	37.27	279,484	71.00	406,164	120.64
Residential Low-Income Sector - Cumulative Projected Portfolio Savings[2]	6,379	0.95	26,642	7.27	47,297	13.68	68,562	23.42
Commercial/Industrial Small Sector - Cumulative Projected Portfolio Savings[2]	27,503	5.33	168,854	36.73	361,698	78.27	617,389	134.60
Commercial/Industrial Large Sector - Cumulative Projected Portfolio Savings[2]	5,669	0.99	33,645	26.90	76,166	55.27	135,311	101.82
Governmental/Non-Profit Sector - Cumulative Projected Portfolio Savings[2]	5,982	1.06	37,506	11.38	79,996	23.72	134,554	42.34
EE&C Plan Total - Cumulative Projected Savings	**75,180**	**12.28**	**419,907**	**119.54**	**844,641**	**241.93**	**1,361,980**	**422.82**
Percent Reduction From Baseline	0.2%	0.2%	1.1%	1.8%	2.2%	3.7%	3.6%	6.4%
Commission Identified Goal			382,144				1,146,431	296.64
Percent Savings Due to Portfolio Above or Below Commission Goal			9.9%				18.8%	42.5%

1. Commission approved Energy Consumption Forecasts and Historical Peak Loads per Energy Consumption and Peak Demand Reduction Targets Order at Docket No. M-2008-2969887, entered March 30, 2009.
2. Adjusted for weather and extraordinary load as applicable.

2.2 EDC Program Lists

According to the EDCs' latest available EE&C Plans, the EDCs plan to implement approximately 150 programs across all customer sectors. The following tables contain a list of the various programs to be implemented by each of the EDCs in the following market sectors: Residential (exclusive of Low-Income), Low-Income, Small Commercial and Industrial, Large Commercial and Industrial, and Governmental/Non-Profit.

Table 2-8: West Penn Power - Program Summaries

West Penn Power
Residential Portfolio Program (exclusive of Low Income)
Compact Fluorescent Lighting (CFL) Rewards Program
Critical Peak Pricing (CPR) Rate
Residential ENERGY STAR® and High Efficiency Appliance Program
Residential Home Performance Program
Programmable Controllable Thermostat (PCT) Program
Residential HVAC Efficiency Program
Residential Efficiency Rewards Rate
Pay Ahead Smart Service Rate
Residential Low-Income Sector Programs
Residential Low Income Home Performance Check-Up Audit & Appliance Replacement Program
Residential Low Income Joint Utility Usage Management Program
Residential Low Income Room Air Conditioner Replacement Measure
Commercial/Industrial Small Portfolio Program
Commercial HVAC Efficiency Program
Commercial Lighting Efficiency Program
Contracted Demand Response Program
Custom Technology Applications Program
Time of Use (TOU) with Critical Peak Pricing Rate
Hourly Pricing Option (HPO) Rate
Commercial/Industrial Large Portfolio Program
Custom Applications Program
Customer Load Response Program
Distributed Generation Program
Commercial and Industrial Drives Program
Governmental/Non-Profit Portfolio Programs
Governmental/Non-Profit Lighting Efficiency Program

Table 2-9 Duquesne Light - Program Summaries

Duquesne Light
Residential Portfolio Program (exclusive of Low Income)
Residential Energy Efficiency
Residential: Schools Energy Pledge
Refrigerator Recycling
Demand Response: Residential A/C Cycling
Residential Low-Income Sector Programs
Low Income Energy Efficiency
Commercial/Industrial Small Portfolio Program
Commercial Umbrella (Small)
Office Buildings (Small)
Retail Segments (Small)
Education (Small)
Industrial Umbrella (Small)
Industrial Mixed (Small)
Demand Response: Small & Mid-Sized Commercial/Industrial
Commercial/Industrial Large Portfolio Program
Commercial Umbrella (Large)
Office Buildings (Large)
Healthcare (Large)
Retail Segments (Large)
Education (Large)
Industrial Umbrella (Large)
Primary Metals (Large)
Chemicals (Large)
Demand Response: Curtailable Load for Large Commercial/Industrial
Governmental/Non-Profit Portfolio Programs
Public Agency (Large)

Table 2-10: Met-Ed - Program Summaries

Met-Ed
Residential Portfolio Program (exclusive of Low Income)
Demand Reduction
Home Energy Audits
Appliance Turn-In
EE HVAC & Solar
EE Products
New Construction
Whole Building Comprehensive
Multiple Family
Residential Low-Income Sector Programs
Warm Plus
Original Warm
Home Energy Audits
Appliance Turn-In
EE Products
Commercial/Industrial Small Portfolio Program
Energy Audit
Equipment Rebate
Multiple Family
Commercial/Industrial Large Portfolio Program
Equipment Rebate
Industrial Motors and VDS
Governmental/Non-Profit Portfolio Programs
Governmental & Institutional

Table 2-11: Penelec - Program Summaries

Penelec
Residential Portfolio Program (exclusive of Low Income)
Demand Reduction
Home Energy Audits
Appliance Turn-In
EE HVAC & Solar
EE Products
New Construction
Whole Building Comprehensive
Multiple Family
Residential Low-Income Sector Programs
Warm Plus
Original Warm
Home Energy Audits
Appliance Turn-In
EE Products
Commercial/Industrial Small Portfolio Program
Energy Audit
Equipment Rebate
Multiple Family
Commercial/Industrial Large Portfolio Program
Equipment Rebate
Industrial Motors and VDS
Governmental/Non-Profit Portfolio Programs
Governmental & Institutional

Table 2-12: Penn Power - Program Summaries

Penn Power
Residential Portfolio Program (exclusive of Low Income)
Demand Reduction
Home Energy Audits
Appliance Turn-In
EE HVAC & Solar
EE Products
New Construction
Whole Building Comprehensive
Multiple Family
Residential Low-Income Sector Programs
Warm Plus
Original Warm
Home Energy Audits
Appliance Turn-In
EE Products
Commercial/Industrial Small Portfolio Program
Energy Audit
Equipment Rebate
Multiple Family
Commercial/Industrial Large Portfolio Program
Equipment Rebate
Industrial Motors and VDS
Governmental/Non-Profit Portfolio Programs
Governmental & Institutional

Table 2-13: PECO - Program Summaries

PECO
Residential Portfolio Program (exclusive of Low Income)
CFL Initiative
Whole Home Performance
Home Energy Incentives
Residential New Construction
Appliance Pickup
Residential Direct Load Control
Residential Super Peak TOU
Residential Low-Income Sector Programs
Low-Income Energy Efficiency
Commercial/Industrial Portfolio Program
Commercial/Industrial Equipment Incentive
Commercial/Industrial New Construction
Commercial/Industrial Direct Load Control
Commercial/Industrial Super Peak TOU
DR Aggregator Contracts
Distributed Energy Resources
Permanent Load Reduction
Governmental/Non-Profit Portfolio Programs
Government/Public/Non-Profit Facility Energy Savings
Programs Common to All Sectors
Conservation Voltage Reduction

Table 2-14: PPL - Program Summaries

PPL
Residential Portfolio Program (exclusive of Low Income)
Appliance Recycling Program
Energy-efficiency Behavior & Education
Energy Assessment & Weatherization Program
Direct Load Control Program
Efficient Equipment Incentive Program
Compact Fluorescent Lighting Campaign
ENERGY STAR® New Homes Program
Time of Use Rates
Renewable Energy Program
Residential Low-Income Sector Programs
E-Power Wise
Direct Load Control Program
Compact Fluorescent Lighting Campaign
Time of Use Rates
Low-Income WRAP
Commercial/Industrial Small Portfolio Program
Commercial and Industrial Custom Incentive Program
Direct Load Control Program
Efficient Equipment Incentive Program
Small Commercial HVAC Tune-up Program
Time of Use Rates
Compact Fluorescent Lighting Campaign
Commercial/Industrial Large Portfolio Program
Load Curtailment Program
Commercial and Industrial Custom Incentive Program
Efficient Equipment Incentive Program
Governmental/Non-Profit Portfolio Programs
Commercial and Industrial Custom Incentive Program
Direct Load Control Program
Efficient Equipment Incentive Program
HVAC Tune-up Program
Time of Use Rates
Renewable Energy Program
Curtailment Program

2.3 Energy Efficiency and Conservation Program Costs and Benefits

The following table contains a summary of the portfolio Benefit-Cost ratios for each of the market sectors by EDC. The ratios were calculated according to the TRC Order.[16]

Table 2-15: Portfolio Summary of Benefit-Cost Ratios by Sector and EDC[17]

	West Penn Power	Duquesne	Met-Ed	Penelec	Penn Power	PECO	PPL
Residential (exclusive of Low Income)	2.4	6.3	2.5	2.1	2.0	2.1	3.3
Residential Low-Income	1.4	4.1	2.1	2.5	2.1	1.7	1.4
Commercial/Industrial Small	5.5	6.8	3.2	2.4	2.3		3.2
Commercial/Industrial Large	3.3	5.9	1.2	1.3	0.9	1.3	2.9
Governmental/Non-Profit	7.0	4.3	2.1	1.7	1.5	1.7	2.4
DR		1.8					
Programs Common to All Sectors						8.2	
Total Portfolio	4.1	5.6	2.4	2.0	1.7	1.8	2.9

The following table contains the portfolio costs per MWh saved for each of the EDCs. This table is based on dividing the filed budget (reflecting the 2% rate cap) by the 3% energy savings target from each of the EDC's filings (i.e., $ available to spend by MWh each EDC must achieve.)

Table 2-16: Summary of EDC Cost per First Year MWh Saved

	West Penn Power	Met-Ed	Penelec	Penn Power	DLC	PECO	PPL
$/Plan MWh saved	$150	$223	$213	$187	$185	$289	$215

[16] The TRC benefit/cost ratios for similar programs across EDCs vary because EDCs have different electric avoided costs, different economic/demographic characteristics, different levels of program participation, different availability of alternative fuels, and different programmatic approaches. Furthermore, for demand response programs, the definition of TRC costs is not consistent across the seven EDCs.

[17] All values are planning estimates per filed plans.

3 Guidance on Program Planning and Reporting Savings

This section describes the responsibilities of the EDCs, which serve as the primary parties responsible for delivering energy and demand savings, during the program planning and implementation stages of Act 129. Key elements include baseline studies, claiming savings, interim measure protocols, custom measure protocols, and reporting results. The following sections will discuss each element in detail.

3.1 Baseline Studies

This section explains the critical importance of understanding baseline energy efficiency levels for energy efficiency and demand response measures or programs. Energy and demand savings cannot be directly measured. Instead, savings are determined by comparing energy use after a program is implemented, which can be measured, with what would have occurred had the program not been implemented, which cannot be measured. This can also be described as the difference in energy consumption between the efficient case and the baseline case, respectively.

The baseline condition and the energy use after the program is implemented can be compared after the data are adjusted so that only program effects are considered when determining energy savings (i.e., data normalized for variations in weather, occupancy, etc.). Given the severity of the 2008 to 2010 economic down turn, particular attention should be paid to the influence of the economic recovery on energy use and subsequent savings associated with conservation programs. For example, equipment utilization will increase as economic activity expands and this may reduce anticipated per unit savings; these changes in utilization or occupancy need to be considered.

To achieve these objectives, the EDCs, with the support of the SWE, will need to:

- Develop and update the descriptions of existing and emerging program markets (e.g., sub-markets, the impact of economic influences and role of market actors)
- Identify market trends, as compared to program participation rates
- Develop baselines and track changes in program penetration, perception and value
- Identify shifts in the primary and sub-markets that would signal market transformation progress over time, with a focus on pertinent market indicators that are likely to increase understanding of program performance.
- Where appropriate, the EDCs should build upon prior evaluation efforts.

The following steps will be taken to develop the baseline studies:

- Identify and prioritize needed baseline studies;
- Review secondary data sources, including reports previously prepared for the Commission and for other organizations, and assess the sources' value for defining baseline conditions in PA; and
- Coordinate evaluation efforts to maximize the efficiency of data collection, research, and reporting efforts, including sharing market data to estimate market effects attributable to an EE&C program's presence in the market.

Baseline definitions consist of site-specific issues as well as policy-related considerations (e.g., state or federal energy standards or codes). The site-specific issues include the characteristics of equipment in place before an efficiency measure is implemented and how and when the affected systems are operated. In many retrofit situations, the baseline condition for a measure can be established when a new energy efficient measure is installed. For example, one can calculate the difference in power consumption (watts per fixture or appliance) and then calculate energy and demand savings for particular installations.

However, it is also important to consider where in the life-cycle of the existing system the new equipment was installed. The options are:

- Early replacement of equipment before the end of its useful life;
- New equipment installed for failed equipment; or
- New construction.

In the case of new construction, it may be desirable to physically audit a sample of new residences and buildings built in recent years. For example, in order to better understand the residential new construction market in Pennsylvania, establish a baseline condition, and build up a base of knowledge to help in strategically approaching builders, a comprehensive baseline study of local residential new construction practices may be necessary.

Per the RFP and roles and responsibilities chart reproduced in Section 1.2, EDCs are required to participate collaboratively in these baselines studies. The SWE will provide support for facilitating such collaborative efforts and reviewing study instruments. Additionally, we recommend that these studies be conducted sooner rather than later, in order to obtain a meaningful representation of the baseline statistics of the state prior to market transformations resulting from the implementation of successful EE&C Plans. After discussions with the PA PUC staff, the EDCs and the SWE Team during the summer of 2011, the scope of work and budget for the SWE Team was expanded to include conducting statewide residential and commercial baseline studies. These studies are expected to be completed by February 1, 2012. The results of these baseline studies will then be used in the statewide market potential study being conducted by the SWE Team. The statewide market potential study is expected to be completed by March 1, 2012.

3.2 Guidance on Calculating Claimed Savings

The purpose of this section is to provide a framework in which all EDC programs with diverse measure offerings and unique implementation strategies can report gross savings in a manner that is equitable and consistent between EDCs. Applying different standards to the same measures in different service territories compromises the PA PUC's ability to objectively determine compliance with the statutory targets. For the purposes of Act 129, reported gross savings are determined by a savings protocol using a method described here. Three types of savings protocols exist that may be used to determine savings:

1. TRM protocols; or standard protocols approved by the Commission and in the TRM.
2. Interim measure protocols; or standard protocols approved by the SWE but not yet formally approved by the Commission.
3. Custom measure protocols; or protocols not suitable for the TRM that involve unique variables and/or whose results are measured directly.

3.2.1 Technical Reference Manual (TRM) Protocols

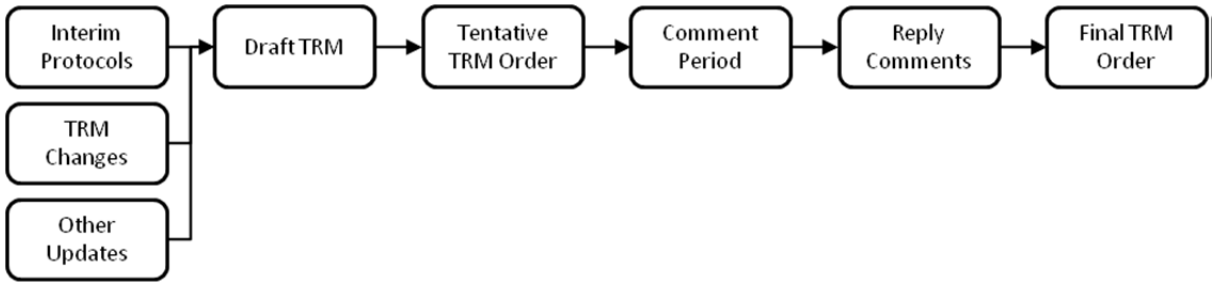
The TRM, updated and adopted each year by the Commission as a component of the Energy Efficiency and Conservation Program evaluation process through the EE&C Program Implementation Order entered on January 16, 2009, contains protocols that determine savings for "standard measures that warrant standard energy efficiency calculation methods and assumptions"[18] by either deeming savings or providing an algorithm with open variables to calculate savings. These protocols, also called TRM protocols, must be used to claim savings. Two exceptions exist:

1. All custom measures, or measures not included in the TRM
2. Alternative measurement methods subject to review and approval by the Commission to ensure their accuracy[19]

These exceptions will be discussed in the Custom Measure Protocol and Interim Measure Protocol sections, respectively.

In order for the TRM to be an effective tool for Act 129, the Commission ordered a regular annual update to the TRM. It is the intention of the Commission to update and expand the TRM, which was "initially created to fulfill requirements of the AEPS Act", on an annual basis to "fulfill the requirements of the EE&C provisions of Act 129".[20] All changes made during the TRM update process will be prospective and thus will not retrospectively affect savings determinations for the program year already underway, unless otherwise determined by the Commission. Updates to the TRM will involve the typical stakeholder process which follows the Tentative Order, Comment Period, Final Order procedure (see Figure 3-1).

Figure 3-1: TRM Update Process (Annual)



[18] 2009 TRM Order, p. 10
[19] 2009 TRM Order, p. 14-15
[20] Implementation Order, p. 13

The estimated schedule for approval of the TRM and effective dates is shown in Table 3-1. Actual dates of deadlines leading up to the approval of the TRM are subject to change by the Commission. The intent of the six month period between the approval date and the effective date is to allow EDCs sufficient time to update program documents, marketing materials, tracking databases, and other miscellaneous program or system processes. The SWE Team notes that the time period between the approval date and the effective date has been less than six months for the 2009, 2010 and 2011 TRMs.

Table 3-1: Timeline for TRM Updates

TRM Version	Program Year	Approval Date	Effective Date[21]
2009 TRM	**PY 1**	June 1, 2009	June 1, 2009 – May 31, 2010
2010 TRM	**PY 2**	June 8, 2010	June 1, 2009 – May 31, 2011[22]
2011 TRM	**PY 3**	February 28, 2011	June 1, 2011 – May 31, 2012
2012 TRM	**PY 4**	December 31, 2011*	June 1, 2012 – May 31, 2013

*****Estimated date

A collaborative TRM Technical Working Group (TWG), comprised of BTUS staff, SWE, EDCs, and EDC teams, has been initiated to review, clarify, improve, and add new savings protocols to the TRM. Generally, the mission of this group is to provide technical guidance to the Commission relative to the quantification of energy and demand savings. After the first year, during which large changes to the TRM were made, changes will be reviewed throughout the year through the interim TRM process (see Section 3.2.2).

Impact evaluation results, as they are made available, will be used to update the TRM during the annual TRM update process. The applicability of these findings will be investigated and discussed through the TWG review process to ensure that the best available estimates of energy and demand savings are presented in the TRM. It is expected that because most of the original TRM protocols are based on secondary research, impact evaluation findings will serve as more accurate representations of the actual savings.

3.2.1.1 TRM Deemed Savings Measures

The TRM includes a large number of measurement protocols that have deemed savings. This type of protocol is particularly appropriate for residential customers that have similar electricity usage characteristics. For measures using deemed savings the TRM does not specify either pre- or post-site verification of the installation quantity and often uses broad classifications of equipment type. This is attractive because of its simplicity and is particularly practical for programs that distribute equipment at the retail level or for "give-away" programs.

[21] Note: The effective date is the period of time during which the TRM is being used.

[22] The 2010 TRM Order was entered on June 8th, 2010 and applied retroactively to the start of the Act 129 Program, per direction of PUC. See 2010 TRM Order, pages 16-17. This effectively overrode all 2009 TRM protocols.

Recommendations of the SWE relative to TRM deemed savings for future years include:

- Maintain an active TWG, chaired by the SWE, including technical experts from the utilities and other independent experts as may be needed to clarify and improve the TRM.
- Conduct a national review of Deemed Savings Databases to determine how others have used this tool and the assumptions they have utilized.
- Note the percentage of overall portfolio savings that may be reasonably deemed, based on national best practices and the differences between the market sectors to which deemed savings approaches are best applied.
- Examine the literature referenced in the TRM, which supports the deemed savings assumptions, including the review of the population or tests from which the data was derived and recommendations as to the appropriate population or technologies to which the generalizations are to be applied in Pennsylvania.
- Continue to update the TRM on an annual basis.

3.2.1.2 TRM Partially Deemed Savings Measures

Many of the commercial and industrial programs sponsored by the utilities utilize savings measurement protocols that are based on partially deemed savings. These programs use TRM protocols which include algorithms with key variables that are stipulated, but do not directly stipulate energy savings. Because of the diversity of equipment and factors effecting utilization in the C&I sectors, the application of these methods across a broad class of C&I customers may have limitations that need interpreting.

The approach envisioned for program year two and beyond, for partially deemed savings protocols, must also incorporate a collaborative process and consider the practical limitations imposed by budget constraints, just as in the deemed savings protocols. It must also consider what is reasonably possible within similar budget constraints in other jurisdictions and where those budget dollars are best spent to maximize the overall accuracy of the evaluation process. For programs with significant anticipated impacts and potential savings variance, it is recommended that greater care be taken to provide the rate payers assurance that the savings are based on a reasonable level of rigor. This may include more careful review of reference literature forming the basis of key assumptions. Potential recommendations from the SWE relative to TRM partially deemed savings for year two include:

- Utilizing the TWG recommended above, the SWE will recommend protocols for high impact measures for immediate review and clarification or modification.
- The SWE will work in collaboration with the EDCs to review each protocol and analyze the specific algorithms and definitions of terms to make sure the protocols use accepted industry standards, reasonably estimate savings and that the methodologies for implementing the protocol are clearly defined and can be effectively implemented in practice.
- As part of the evaluation framework, the SWE will examine the literature referenced in the TRM supporting key variables used in partially deemed savings algorithms for discussion in the TWG. This will include reviewing the population from which source data was derived and providing

recommendations as to the appropriate population or technologies to which the generalizations are to be applied.

3.2.1.3 Custom Measures

Custom measures are referenced in the TRM, but they "are considered too complex or unique to be included in the list of standard measures provided in the TRM."[23] Rather than following TRM protocols, savings for custom measures are determined by following custom measure protocols, which follow a different process than TRM protocols. See section 3.2.3 for additional detail.

3.2.2 Interim Measure Protocols

Interim Measure Protocols are the second type of savings protocols that are followed to claim energy savings. The TRM serves as a Commission approved resource containing savings protocols to be used to determine, claim and verify savings that count toward the statutory savings targets for the purposes of Act 129. However, the 2009 TRM Order did not contain protocols for every measure offered through the EDC programs. Given the research required by the TWG to develop sufficient TRM protocols, the processes required by the Commission to approve an order, and the time required by EDCs to update programs consistent with new TRM updates, a vast number of measures were not adopted in the 2010 TRM Order.

The Commission recognized that it was necessary to expand the TRM by adding more savings protocols from the early stages of Act 129.[24] The 2009 TRM Order directed that the TRM be updated on an annual basis, which would allow new savings protocols to be added to the TRM each year.[25] Nevertheless, parties expressed concern that without approved protocols, EDCs would hesitate to offer additional measures that have potential to cost-effectively achieve energy and demand savings.

In order to address this concern, interim protocols were introduced. Interim protocols fit within the boundaries of a TRM protocol (i.e.., standard measures that warrant standard energy efficiency calculation methods and assumptions) but have not yet been formally included in the TRM. Interim protocols exist primarily to allow EDCs to claim savings for measures that do not have TRM protocols, and also to approve alternative measurement methodologies where the TRM is insufficient.

3.2.2.1 Interim Protocol Approval Process

The interim measure protocol approval process has been outlined by Guidance Memo #008. Please see the SWE SharePoint site for the latest version.

[23] Implementation of the Alternative Energy Portfolio Standards Act of 2004: Standards for the Participation of Demand Side Management Resources – Technical Reference Manual Update. Docket No. M-00051865. Annex: Technical Reference Manual (TRM) for Pennsylvania Act 129 Energy Efficiency and Conservation Program and Act 213 Alternative; Page 7.
[24] 2009 TRM Order, Page 10.
[25] Ibid. Page 17-18.

3.2.3 Custom Measure Protocols (CMPs)

Custom measures are defined as measures for which savings calculation methods are not described in the Technical Reference Manual (TRM). Generally, complex measures requiring metering, logging, modeling, or billing analysis to determine impacts are excluded from the TRM and are therefore custom measures. Whereas TRM measures are reviewed and approved by the PUC through the TRM update process, custom measures do not undergo the same approval process.

Since CMPs are outside of the TRM, no specific protocols or methods are set forth to quantify savings. The SWE Team recommends that these protocols be established in general conformity to the IPMVP or Federal Energy Management Program M&V Guidelines. The SWE, in collaboration with the EDCs, will help define general guidelines for Custom Measures that have the potential to capture significant savings. This will serve to limit the risk that utilities have in implementing Custom Measures and provide the utilities with the general expectations for M&V.

EDCs are not required to submit CMPs to the SWE for each measure/technology type prior to implementing the custom measure. EDC implementers and EDC evaluators are not subject to requirements set forth in previously approved CMPs[26,27].

EDCs may report ex-ante savings for a custom measure according to methodologies used by the customers or contractors and approved by the EDC implementers. The measurement methodology used by the EDC implementers is subject to guidelines developed by the SWE[28]. EDC implementers will be required to perform measurements consistent with IPMVP options to collect baseline and/or post retrofit information for custom measures that have estimated savings above a threshold kWh level[29]. They are not limited from performing measurements for custom measures with estimated savings below the threshold. The verification of savings for custom measures is discussed in Section 4.1.2. As part of its role to perform EM&V services for the Commission, the SWE reserves the right to audit and review claimed and verified impacts of all custom measures.

For further information regarding the custom measure process, please refer to Guidance Memo #008. Please see the Share Point site for the latest version.

[26] As an exception, EDCs are required to adhere to previously approved CMPs for the following: Conservation Voltage Reduction (CVR), Low-Income, O Power, Schools for Energy Efficiency. The CVR CMP was updated on September 27th, 2011 in draft form.

[27] This does not limit evaluators from using CMPs approved under the previous process, but allows evaluators to use alternate methods as needed. EDC implementers and evaluators are encouraged to use the methodologies contained in those CMPs as deemed appropriate.

[28] Guidelines will not be burdensome and will be in line with standard program implementation practices. The guidelines will be issued by the SWE on behalf of the Commission and CEEP staff and will ensure adequate transparency and standardization in the ex-ante savings estimation. Guidelines will be added to the Act 129 Audit Plan update.

[29] The threshold kWh is recommended by the SWE to separate large custom projects from the small custom projects. The threshold should be developed by the EDCs and proposed to the SWE for approval. Projects below the threshold may include a baseline assessment, depending on the type of measure. The threshold may vary for different project types.

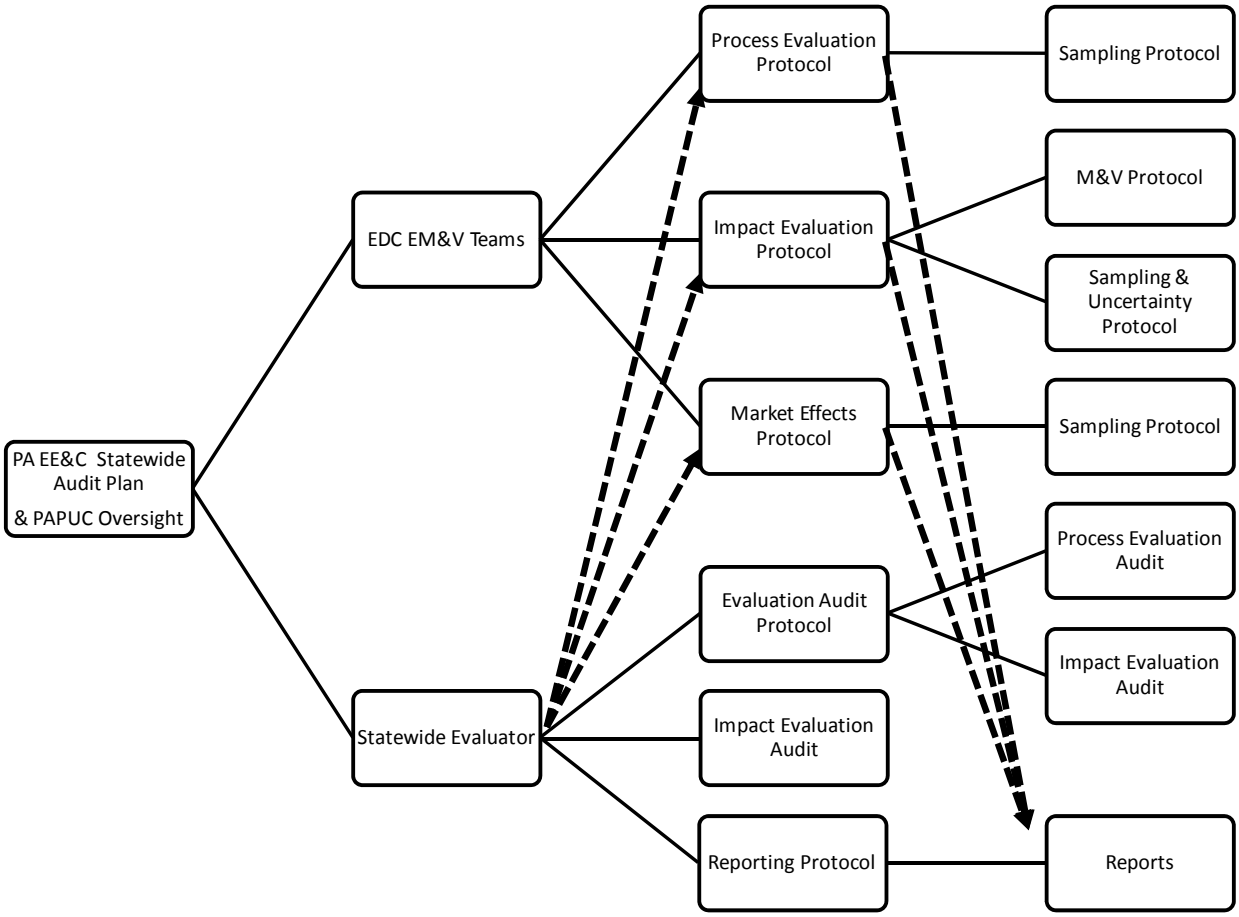
3.2.4 Reporting Savings

Issues related to reporting savings have been outlined by Guidance Memo #006. Please see the SWE Share Point site for the latest version.

4 Guidance on Evaluation Measurement and Verification (EM&V)

There are many components, moving pieces, and concurrent developments involved in the evaluation process of a set of energy efficiency and demand response programs. For the EDC EE&C programs, it is no different. A diagram of the evaluation and audit protocols is depicted in Figure 4-1. The solid lines indicate an action (e.g., the EDCs are responsible for developing the actual Impact Evaluation Protocols to be used in the EM&V process) while the dotted lines indicate influences (e.g., the SWE has outlined certain expectations of the Impact Evaluation Protocols which are to be specified by the EDCs in their EM&V Plans). The components of this figure include roles, responsibilities, guidelines, expectations and requirements of the entire evaluation process, which will be discussed in greater detail throughout this Audit Plan.

Figure 4-1: Operational Overview of Evaluations, Audits and Protocols



4.1 Guidance on Impact Evaluation

Evaluation is the process of independently determining and documenting the results, benefits, and lessons learned from an energy efficiency program. Evaluation results can be used in planning future programs and determining the value and potential of a portfolio of energy efficiency programs in an integrated resource planning process. The results can also be used to assess the performance of implementers responsible for implementing efficiency programs, the effectiveness of program incentives, and the effectiveness of motivators such as payments and penalties.

Impact evaluations determine program-specific induced benefits, which include reductions in energy and demand usage (such as kWh, kW, and therms) and avoided air emissions[30] that can be directly attributed to an energy efficiency program. There are many stages to an impact evaluation. At each stage of the evaluation process, there are typical decisions that must be made based on the desired accuracy and certainty of the evaluation, the funds available, and the anticipated impact and its uncertainty. Table 4-2 outlines the general stages and protocols of the impact evaluation process. The stages will be discussed in greater detail later in this section of the Audit Plan.

Gross impact evaluations will be conducted by the EDC evaluators to determine directly achieved program benefits. The EDC evaluators will utilize data collected during program implementation or conduct independent data gathering activities if the data collected by program implementers is unreliable. The M&V activities may include surveys or direct observation and measurement of equipment performance and operation at a sample of participant sites to verify that the energy savings reported for the participant at the program level is correct. Successful impact evaluations assess the costs incurred with the Value of Information (VOI) received, i.e. they appropriately balance the level of evaluation detail (rigor) to the level of effort required (cost). How deeply an evaluator goes into the measurement of key variables (rigor) at a sampled participant's site depends on the value of that information in confirming claimed savings.

There is a hierarchy in the process of developing data for the evaluation of the EDC programs. The savings protocols for efficiency measures define how the savings will generally be calculated by EDC program administrators. The impact evaluation protocols are the procedures followed by the third party independent EDC evaluators to verify the energy and demand savings claimed by the EDC implementers. The SWE audit protocols are the protocols followed by the SWE to audit and confirm the EDC evaluators' verified impacts.

[30] Air emissions are not currently included as part of the EE&C program impact evaluations; however, estimates of emission reductions can easily be estimated using Energy Information Administration and Federal Energy Regulatory Commission reports and gross energy savings.

Table 4-1: Impact Evaluation Hierarchy

Level	Description	Requirements/ Discussion Points
1. Program Implementation • Deemed Savings • Partial Deemed • Custom/Unspecified	M&V Protocols and Site Specific M&V plans used to calculate claimed savings • TRM protocols • Interim TRM protocols • Custom measure protocols	1. TRM protocols are used for prescriptive measures 2. Implementers approve M&V protocols for custom measures 3. Savings are claimed using the approved protocol
2. EDC Impact Evaluation	EDC evaluator samples program data and calculates Realization Rate with approved EM&V protocols	1. Statistical Sample of participants analyzed 2. Field engineering and site-specific analysis 3. Calculation of Realization Rates
3. SWE Audited Impact	SWE works with EDC evaluators to audit and ensure accuracy of reported savings and Realization Rate or conducts independent sample if needed	1. Collaboration with EDC Impact Evaluation activities (i.e., joint site-visits) 2. Independent site visits and field verification 3. Recommendations to adjust Realization Rates

4.1.1 Impact Evaluation Research Objectives

The following sections outline the evaluation framework expectations and guidelines necessary to address the following research objectives:

- Determine Method for Calculating Savings.
- Set acceptable levels of Rigor, Precision and Bias for M&V activities.
- Determine Realization Rates for Verified Gross Savings.
- Determine Net to Gross (NTG) Ratios for Net Savings.

The first three research objectives relate to verified gross savings and are addressed in Section 4.1.2. The fourth research objective relates to net savings and is addressed in Section 4.1.3.

4.1.2 Calculating Verified Gross Savings

One of the primary research objectives of an impact evaluation is to calculate the verified gross savings, which are the savings directly achieved by program benefits validated by an independent third party evaluator. Impact evaluations begin at the installation or site level. EDC evaluators verify savings according to the appropriate savings protocols relative to the claimed savings for a statistically

significant number of sites and calculate a realization rate. This realization rate is then applied to the population at large to determine the verified gross savings. Where applicable, the results of these EDC impact evaluations will also be used to update the kWh and kW savings in the Pennsylvania PUC Technical Reference Manual so that the Manual reflects the latest available information on measure and program savings. The following sections provide detailed guidance for EDC evaluators for calculating verified gross savings for impact evaluations.

4.1.2.1 Determine the Measure Type

Much of the savings anticipated by the Act 129 programs will be measured through methods set forth in the TRM. Though metering studies and detailed analysis are encouraged to inform future updates of savings protocols in the TRM, measures with TRM protocols must be verified by the EDC evaluators using TRM protocols and assumptions. A large portion of these savings are either (1) deemed based on units installed, sold or given away or (2) subject to generic assumptions relative to the performance of the technologies and how the technologies are used. In many cases, applying the assumptions in the TRM by definition "verifies" the savings if the savings measure is installed. The methods identified in the TRM for quantifying savings fall into three general categories shown below in Table 4-2.

Table 4-2: Measure Categories to be Verified

Categories	Description	Examples	Assumptions	Quantification Required
TRM Deemed Savings Measures	Deemed kW and kWh	CFL's Appliances	Baseline, Hrs, Impact Baseline, Hrs, Impact	Number, Type Number, Type
TRM Partially Deemed Measures	Savings Algorithms with Open Variables	C&I Lighting C&I Motor	Baseline, Impact Baseline, Impact	Number, Type, Hrs HP, Load Factor
Custom Measures[31]	All Other, Unspecified	Large Lighting Other VFD EMS, Controls		Defined by SSMVP Defined by SSMVP Defined by SSMVP

The measure category dictates the specified M&V activities required to verify that particular measure and are defined as follows:

- For TRM Deemed Measures, site specific data collection will include the number of installations (lighting fixtures, appliances, VFDs, HVAC units etc.) and other site specific inputs as determined on a case-by-case basis. This is at a basic level of rigor, or less, given that the savings are deemed in the TRM.
- The TRM Partially Deemed Measures will follow either a basic or an enhanced level of rigor depending on the type of measure (uncertainty and complexity) and the savings threshold (level

[31] SSMVP may specify the use of sampling for large quantities of measures (i.e. large lighting projects).

of impact). The data collected, such as hours of use, will then be used according to the TRM to estimate the savings.

- The Custom Measures will be dependent on the Custom Measure protocol and will require an appropriate level of rigor to implement the protocol.[32]

In developing evaluation framework guidelines, it is necessary to consider the measure categories currently delineated by the TRM as illustrated in Table 4-2. The methods required to verify and evaluate saving for these different measure categories need to be considered separately, and in each case, form the foundation of any evaluation process designed to verify those specific savings. To the extent that methodologies using these protocols capture the savings that occur as a result of the expenditure of rate payer funds, the evaluation process rests on a sound foundation.

The approach to verifying savings should be clear, technically sound and based on accepted industry standards. The quantification of savings is both an art and a science, as energy savings are the difference between energy that would have been used and what was actually used. In practice, engineering, empirical science and reasonable assumptions need to be used to determine what "would have been used".

4.1.2.2 Determine the Level of Engineering Rigor

The impact evaluation protocol establishes the measurement and verification (M&V) methods and the assigned level of engineering rigor appropriate for the EDC evaluators.[33] Level of engineering rigor is defined as the level of detail involved in the verification of the EDC reported impacts and defines the minimum allowable methods to be used by the EDC evaluators to verify the savings claimed by the EDCs. In this way, the impact evaluation protocol establishes a minimum level of detail in order to ensure that the verified gross savings are at the level of reliability needed to support the overall reliability of the savings in reference to statutory savings targets. The impact evaluation protocol provides guidelines on the evaluation methods used by the EDC evaluators for specific evaluation groups. These groupings consist of multiple programs (program components/measures) having common characteristics that provide evaluation efficiencies in the contracting, supervision and implementation of evaluation efforts. The groupings will typically include similar types of programs or measures (e.g., residential appliance rebates, commercial rebates, commercial lighting, etc.).

The evaluation protocols define two levels of rigor, ***basic*** and ***enhanced***. Each level of rigor provides a class of minimum allowable M&V methods, which are based on standard evaluation practices, in order to offer flexibility for the EDC evaluators to assess and propose the most accurate and cost effective

[32] Please refer to GM-002 for further instructions regarding custom measure protocols.

[33] The following definitions are important for the reader to understand:
Impact Evaluation - An evaluation of the program-specific directly induced quantitative changes (e.g. kWh, kW, and therms) attributable to an energy efficiency program.
Measurement and Verification (M&V) - A subset of program impact evaluation that is associated with the documentation of energy savings at individual sites or projects using one or more methods that can involve measurements, engineering calculations, statistical analyses, and/or computer simulation modeling.

methods to verify gross savings, by balancing cost and rigor. The choice of basic rigor versus enhanced rigor will depend on the type of measure, relative complexity of savings calculations, the level of uncertainty, and most importantly, savings impact. Generally, flexibility is allowed to choose the appropriate level of basic or enhanced rigor, however the SWE reserves the right to challenge the level of rigor used by the EDC evaluators and request revision of the verification technique where necessary.

Table 4-3 provides guidelines as to the actions associated with basic rigor and enhanced rigor for specific measure categories. EDC evaluators should follow a method that adheres to the current TRM. As previously stated, these are the *minimum* allowable methods. Evaluators are highly encouraged to take advantage of opportunities to collect additional data that will be useful for future TRM updates that result in more accurate and reliable savings protocols.

The M&V options defined under each level of rigor provide independent evaluators cost-effective methods to verify program impacts without compromising the accuracy of the reviews. In general, the TRM Deemed measures will follow a Basic level of rigor and Custom measures will follow an Enhanced level of rigor. The TRM Partially Deemed Measures will follow either a Basic or Enhanced level of rigor depending on the type of measure and level of impact. These paths are depicted in Figure 4-2, which provides guidance on choosing the level of rigor by measure type.

Figure 4-2: Required Protocols for Impact Evaluations

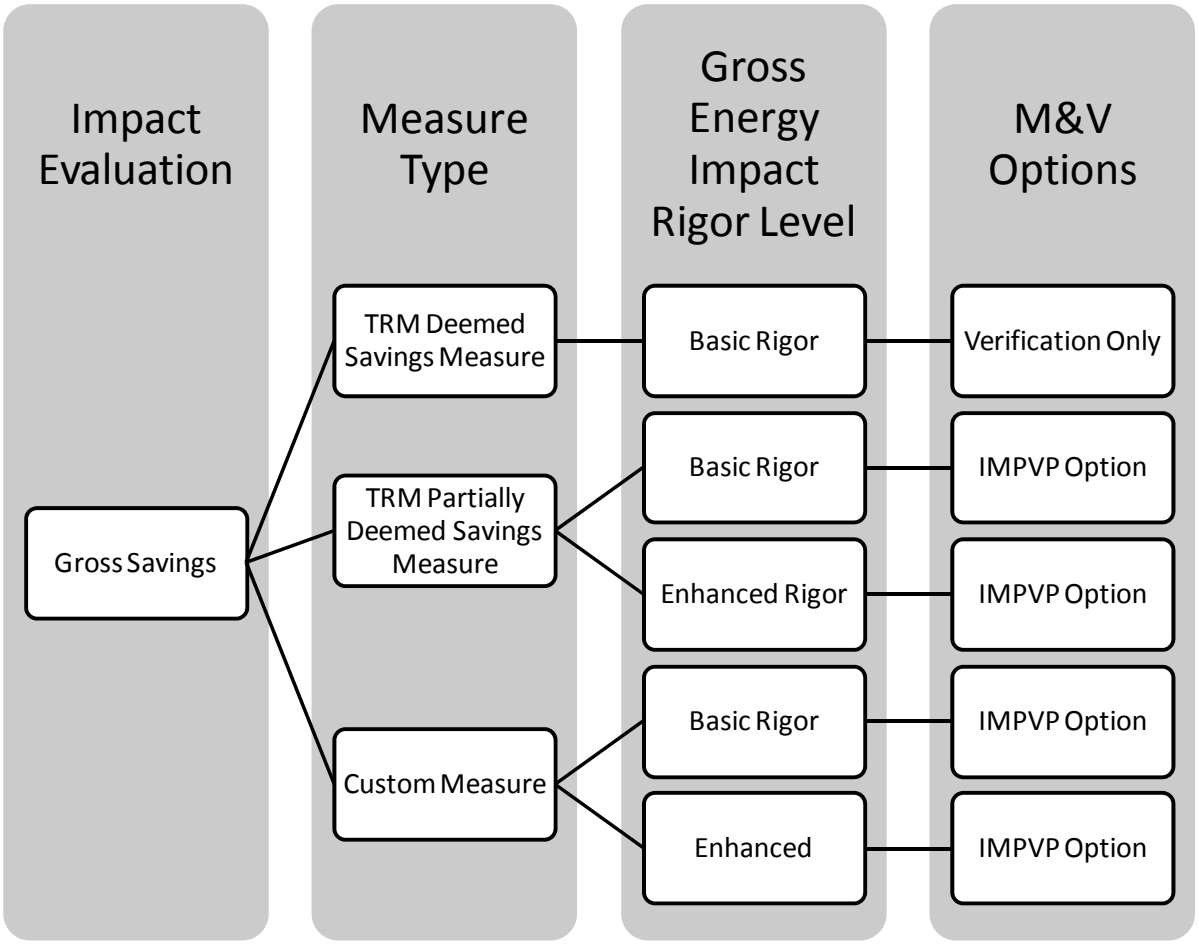


Table 4-3: Definitions of Levels of Rigor for Impact Evaluation

Rigor Level	Minimum Allowable Methods for Gross Impact Evaluation
Basic	1. Verification only analysis for TRM Deemed Measures - verification of number of installations, stipulated operating hours and other assumptions and inputs to the deemed savings estimates specified in the TRM. 2. Simple Engineering Methods with M&V equal to International Performance Measurement and Verification Protocol (IPMVP) Option A for TRM Partially Deemed Measures. Verification of appropriate application of the TRM savings algorithms. Spot measurements and other site specific stipulations where subscribed by the TRM.

Enhanced	1. Retrofit Isolation Engineering methods as described in IPMVP Option B for the TRM Partially Deemed Measures. 2. A fully specified regression analysis (IPMVP Option C) of consumption information from utility bills with adjustments for weather and time of period analysis. The SWE Team recommends that at least twelve (12) months of pre and post retrofit consumption information be required unless program design does not allow for pre retrofit billing data, such as new construction. In these cases well matched control groups and post retrofit consumption analysis is allowable. (*Note: In certain cases either "match groups" or simulation may, in fact, be easier to use. This will be determined on a case-by-case basis.*) 3. Building energy simulation models as described in IPMVP Option D.

4.1.2.2.1 Energy – Basic Rigor Option 1: Verification Only Analysis

The first class of allowable methods for basic rigor is a verification only analysis. This applies mainly to the TRM Deemed Measures. The objective is to confirm that measures are actually installed, operational and installation meets standard requirements. Installation verification shall be conducted at a random sample of sites claiming energy savings. Sampling of measures within a project and sampling at the program level for evaluation purposes will be specified according to the Sampling and Uncertainty Protocols.

4.1.2.2.2 Energy – Basic Rigor Option 2: Simple Engineering Model (SEM)

The second class of allowable methods for basic rigor is a simple engineering model (SEM). An SEM is equivalent to IPMVP Option A. The IPMVP provides overall guidelines on M&V methods; however, more program- or technology-specific guidelines are required for the EDC programs. Section 4.1.2.4 provides guidelines for developing site specific M&V protocols.

SEM are straightforward algorithms for calculating energy impacts for non-weather dependent measures such as energy-efficient lighting, appliances, motors and cooking equipment. These algorithms are defined in the TRM as Partially Deemed Measures. Several algorithms have open variables and require additional stipulations or measurements. Further clarification will be provided by the TWG on the use and application of the TRM algorithms and methods.

Exceptions to this requirement are Custom Measures, which are more complex and typically exhibit significant measure interactions. Evaluations of these measures or programs offering these measures must use the methods falling under the Enhanced Level of Rigor.

Sampling for the M&V used in the SEM must be conducted as prescribed in the Sampling and Uncertainty Protocol (see Section 4.4). In both the evaluation plan and the evaluation report, the inputs selected and the methods selected for the verification/measurement must be justified in terms of why

they are the factors that provide the most likely unbiased and reliable gross energy impact estimates for the evaluation study being conducted.

4.1.2.2.3 Energy – Enhanced Rigor Option 1: Retrofit Isolation Engineering Models

The first class of allowable methods for enhanced rigor is the retrofit isolation measurements, as described in Option B of the IPMVP. This method is used in cases where full field measurement of all parameters for the energy use for the system in which the efficiency measure was applied are feasible and can provide the most reliable results in an efficient and cost-effective evaluation. Applying a VFD to a constant speed pump in a variable flow pumping application would be a typical example of when this method would likely be used.

Sampling must be conducted according to the Sampling and Uncertainty Protocol (see Section 4.4).

4.1.2.2.4 Energy – Enhanced Rigor Option 2: Whole Building Simulation

The second class of allowable methods for enhanced rigor is building energy simulation programs calibrated as described in the Option D requirements in the IPMVP. The engineering models that meet the Option D requirements are generally building energy simulation models. This method can be applicable to many types of programs that influence commercial, institutional, residential and other buildings where the measures impact the heating, ventilation or air conditioning (HVAC) end-use. This method is often used for new construction programs and building HVAC or shell upgrades in commercial and residential programs.

In addition, industrial efforts can include changes in process operations where the appropriate type of model could be a process-engineering model. These are specialized engineering models and may require specific software to conduct engineering analysis for industry-specific industrial processes. Where these types of models are more appropriate, the gross energy impact protocol allows the use of a process engineering model with calibration as described in the M&V protocols to meet the enhanced rigor level.

Sampling must be conducted according to the Sampling and Uncertainty Protocol (see Section 4.4).

4.1.2.2.5 Energy – Enhanced Rigor Option 3: Billing Regression Analysis

The third class of allowable methods for enhanced rigor is a regression analysis of consumption data that statistically adjusts for key variables which change over time and are potentially correlated with gross or net energy savings. As a way of capturing the influence of weather, evaluators may incorporate weather-normalized consumption as the dependent variable, or include heating- and cooling-degree days directly in the model. Other variables that change over time that are often correlated with gross and net energy savings include, among others; the state of the economy (recession, recovery, economic growth), fuel prices, occupancy changes, behavior changes (set-points, schedules, usage frequency), changes in operation and changes in schedule. The EDC evaluators are free to select the most appropriate additional variables to include.

Regression-based methods must also meet the requirements of the Sampling and Uncertainty Protocol. Many of the requirements in the Sampling and Uncertainty Protocol require specific actions and

documentation regarding data cleaning, model specification, testing and reporting for regression-based methods.

4.1.2.2.6 Demand – Basic Rigor

The basic rigor level for the gross demand impact protocol prescribes that, at a minimum, on-peak demand savings are estimated based on allocation of gross energy savings through the use of allocation factors, coincidence factors or end-use savings load shapes. These secondary data are available in the TRM. The basic level of rigor is applicable only to the TRM Deemed Measures and Partially Deemed Measures below a threshold specified by the TWG. Custom Measures will follow an enhanced approach.

4.1.2.2.7 Demand – Enhanced Rigor

The enhanced rigor level for the gross demand impact protocol requires primary data from the program participants. This could be interval-metered data, TOU consumption billing data, from field measurement or from billing demand data. Estimation of peak demand savings estimates is required. If the methodology and data used can readily provide 8,760-hour output, these should be provided.

The evaluation of demand impact may employ different methods based on the type of program used to drive these demand savings:

1. Permanent Demand Reduction Measures. These measures will have different peak demand coincidence factors based on both the technology implemented and the time frame targeted for defining a demand reduction.
2. Demand Response Programs. These programs often rely on protocols established by the pertinent Independent System Operator (ISO). Since these resources need to target the 100 highest peak load hours that are dynamic and vary with system peak loads and weather, dispatching this resource is complex. The actual system peak cannot be known in advance and may vary from year to year requiring the EDCs to over shoot their goal in order to be assured of making their goal. This may be costly and unproductive without some rolling average methodology, which the SWE may be able to support as a verification approach with Commission approval.
3. Tariff and Demand Pricing Programs. By providing true, marginal price signals for demand, some EDCs will encourage economic curtailment. Evaluation of these programs will require careful consideration of the appropriate experimental design for the establishment of a control group.

For permanent demand reduction measures, interval meter data, if available, should be used to construct pre- and post-peak hour load shapes. The data should be adjusted for weather, day type and other pertinent variables. If interval meter data is not available, spot or continuous metering/measurement at peak pre- and post-retrofit should be conducted during the 100 highest peak load hours of the utility system. These data will be used with one of two engineering modeling approaches: (1) full measurement IPMVP Option B or (2) calibrated engineering model Option D, where the modeling approach must meet all requirements as provided in the EM&V protocol. Both of these

engineering methods need to be designed to achieve 10 percent precision at a 90 percent confidence level and must meet the requirements of the Sampling and Uncertainty Protocol.

For demand response measures, demand reduction should be determined and verified using the PJM Interconnection Demand Response Protocols[34] where applicable. Hourly peak load reductions from demand response (DR) measures for Direct Load Control (DLC) and Load Curtailment (LC) will be determined in accordance with PJM measurement & verification protocols, related business rules, protocol approval processes and settlement clearing due diligence practices[35] that will be in place during the 2012 summer period (June 1 - September 30, 2012), as verified by the EDC and reviewed by both the EDCs' independent evaluators and the SWE. Peak load reductions from critical peak pricing (CPP) programs will be determined consistent with EDC EM&V Plans and consistent with PJM Customer Baseline methods and business rules, as they may be reasonably applied to the CPP programs. Peak load reductions from DLC, CPP and LC will be determined for each Act 129 DR event hour for June 1, 2012, through September 30, 2012. When determining customer baselines, Act 129 DR event days and PJM DR event days (e.g., for PJM emergencies and economic events for which participants have settlements) will be excluded. For further details, please see the Demand Response section of the TRM.[36]

The SWE audit activities will include verifying compliance to the protocol. The EDC evaluators have a certain degree of flexibility in choosing the class of methods specified in this protocol. The protocol assumes that over 95% of the utility peak hours/event hours will coincide with the PJM event hours, and the EDCs will be allowed to claim savings for the PJM event hours. Further clarification will be provided by the TWG on the aggregation of the EDC system peak demand reduction estimates calculated and reported for the EDC specific 100 peak hours. The aggregation method will be based on a load shape normalization method using a multi variable regression analysis based on independent variables such as total load, weather determinants, etc.

4.1.2.3 M&V Activities

This section provides a list of M&V activities that are acceptable for evaluation activities.

4.1.2.3.1 Existing Equipment

Where applicable and appropriate, it will be necessary to verify the existing equipment and possibly gather the equipment baseline data in order to compute the deemed, partially deemed, and custom savings estimates. The objective would be to verify that the existing equipment is applicable to the program under which it is being replaced. Additionally, the baseline consumption and run-time patterns may be required to complete the engineering calculations used to estimate savings.

[34] See < http://www.pjm.com/~/media/documents/manuals/m18b.ashx> for more information.
[35] See the Secretarial Letter issued by the Commission on January 12, 2011, at Docket No. M-2008-2069887.
[36] Details for demand response are provided starting in the 2012 TRM.

4.1.2.3.2 Measure Installation Verification

The objectives of measure installation verification are to confirm that the measures were actually installed, the installation meets reasonable quality standards, and the measures are operating correctly and have the potential to generate the predicted savings. Installation verification should be conducted at some or all project-specific installation sites claiming energy or peak demand impacts where any M&V activities are conducted. Installation verification activities may also be specified by the Process Evaluation Protocols discussed in Section 4.2.

4.1.2.3.3 Measure Type

Measure existence should be verified through on-site inspections of facilities. Measure, make and model number data should be collected and compared to participant program records as applicable. Sampling may be employed at large facilities with numerous measure installations. As-built construction documents may be used to verify measures, such as wall insulation, where access is difficult or impossible. Spot measurements may be used to supplement visual inspections, such as solar transmission measurements and low-e coating detection instruments, to verify the optical properties of windows and glazing systems.

4.1.2.3.4 Installation Quality Relative to Savings

Measure installation inspections serve two purposes. The first is to ensure that the measure was, in fact, installed. The second purpose is to verify that the new measure is capable of producing the savings predicted.

4.1.2.3.5 Correct Operation and Potential to Generate Savings

Correct measure application and measure operation should be observed and compared to project design intent. For example, for C&I CFL applications in seldom used areas or occupancy sensors in spaces with frequent occupancy should be noted during measure verification activities and modifications to hours of use categories should be made. At enhanced rigor sites,[37] commissioning reports (as applicable) should be obtained and reviewed to verify proper operation of installed systems. If measures have not been commissioned, measure design intent should be established from program records and/or construction documents. Functional performance testing should be conducted to verify operation of systems in accordance with design intent.

4.1.2.3.6 Verification Activities by Measure Type

Examples of these activities applied to specific measure types can be found in Table 4-4.

[37] Sites include Custom Measures and Partially Deemed Measures with high anticipated impact.

Table 4-4: Verification Activities by Measure Type

Measure Type:	Basic Rigor Level:	Enhanced Rigor Level:
Appliances	• Verification of TRM inputs (type of unit, energy source, usage, location)	• Verification of TRM inputs. • Spot measurements (kW). • Short term metering (kW, operating hours).
CFL Rewards/Give Away	• Verification of quantity based on invoices for bulbs purchased by category (wattage, size etc.). • Predefined operating hours based on TRM.	• NA
Weatherization, Envelope Improvements	• Verification of measure installation. • Software simulation for verifying energy savings.	• Verification of measure installation. • Software simulation for verifying energy savings. • Billing Analysis
Residential HVAC Efficiency	• Verification of measure installation (quantity, type, efficiency). • Baseline efficiency defined by TRM (baseline efficiency equals efficiency of old equipment for early replacement; for end of life replacement and new construction baseline efficiency equals efficiency of standard equipment compliant with code). • New equipment efficiency from manufacturers catalog data. • Stipulated operating hours (TRM defined, defined by baseline studies or customer reported)	• Verification of measure installation (quantity, type, efficiency). • Pre and post installation site visits to verify efficiency levels. • Baseline efficiency equals efficiency of old equipment for early replacement; for end of life replacement and new construction baseline efficiency equals efficiency of standard equipment compliant with code). • Short term metering (pre or post) to calculate EFLH.
Residential Lighting	• Verification of measure installation (fixture quantity, type). • Stipulated operating hours (TRM defined)	• NA

C&I Lighting	• Verification of measure installation (fixture quantity, type). • Pre and post fixture types and performance. • Operating hours (TRM defined, undefined).	• Verification of measure installation (fixture quantity, type). • Pre and post fixture types and performance. • Short term metering to log operating hours and stipulated categories.
C&I HVAC Efficiency	• Verification of measure installation (quantity, type, efficiency). • Baseline efficiency defined by TRM (baseline efficiency equals efficiency of old equipment for early replacement; for end of life replacement and new construction baseline efficiency equals efficiency of standard equipment compliant with code). • New equipment efficiency from manufacturers catalog data. • Stipulated operating hours (TRM defined, defined by baseline studies or customer reported)	• Verification of measure installation (quantity, type, efficiency). • Pre (where applicable) and post installation site visits to verify baseline and retrofit equipment information. • Short term or continuous metering (kW) for a minimum of three weeks to calculate pre and post energy use.

4.1.2.4 M&V Data Requirements

For each measure type and level of rigor, data is required for verification. The M&V data requirements can be found in Appendix D for the following programs:

Commercial and Industrial Sector

- Lighting
- Electric Motors
- HVAC
- Variable Frequency Drives

Residential Sector

- HVAC
- New Construction
- ENERGY STAR Appliance
- Refrigerator/Freezer Retirement
- ENERGY STAR Lighting
- ENERGY STAR Windows
- Direct Load Control

4.1.2.5 Site Specific M&V Plan (SSMVP)

A SSMVP is designed to specify the data collection techniques for physical evidence from field installations of energy efficiency technologies. The EDC evaluators must design and document SSMVPs for each measure with a defined savings protocol in the TRM that requires a physical on-site inspection[38]. SSMVPs define the quantitative data that needs to be collected from the field or other primary sources. For the purposes of this Audit Plan, the SSMVP should cover all field activities dedicated to collecting site-specific information necessary to calculate savings according to the engineering equations specified at the participant level and to prepare for an evaluation audit of gross savings impacts. These activities typically include:

- Measure counts
- Observations of field conditions
- Building occupant or operator interviews
- Measurements of parameters
- Metering and monitoring

As part of the process of developing this Audit Plan, the SWE reviewed the TRM protocols and has provided guidance for the SSMVPs relative to most of the protocols set forth in the TRM (see Appendix D). The SSMVPs recommended not only provide guidance as to the information required to perform the calculation in the protocol, but also provide information needed to conduct a credible audit.

4.1.2.5.1 SSMVP for Custom Measures

This section applies to the development of SSMVPs for custom measures.

Field measurements are an important component of determining savings for complex projects. While the IPMVP is written to allow flexibility, its application requires a thorough knowledge of measure performance characteristics and data acquisition techniques. Energy use varies widely based on the facility type and the electrical and mechanical infrastructure in the facility or system. A measurement strategy that is simple and cheap in one building (such as measuring lighting energy at a main panel) may be much more expensive in a similar building that is wired differently. For this reason, EM&V resources, costs and benefits must be called upon and allocated considering the type of measure and program impact.

A SSMVP is required to define how a savings protocol will be implemented in practice. This includes specifying data to be gathered and stored for each field measurement project that documents the project procedures and rationale. This is to ensure that the results can be audited for accuracy and repeatability. Within the guidelines established by the IPMVP, there is considerable latitude for the EDC evaluator to develop a SSMVP and implement the plan in the field. EDC evaluators shall evaluate the uncertainty in the desired data product and develop a SSMVP for a sampled custom measure/project

[38] SSMVPs for TRM measures may contain simplified information and/or be a generic M&V plan. Guidance for custom measures will be developed in conjunction with the EDCs.

that manages the uncertainty in the most cost-effective manner. The contribution of specific engineering parameters to the overall uncertainty in the savings calculations should be identified and used to guide the development of the SSMVP.

The SSMVP for custom measures must include the following sections:

1. Goals and Objectives
2. Building Characteristics and Measure Description
3. M&V Method
4. Data Analysis Procedures and Algorithms
5. Field Monitoring Data Points
6. Data Product Accuracy
7. Verification and Quality Assurance Procedures
8. Recording and Data Exchange Format

The content of each of these sections is described below.

Identify Goals and Objectives: The goals and objectives of the M&V should be stated explicitly in the M&V plan.

Specify Site Characteristics: Site characteristics should be documented in the plan to help future users of the data understand the context of the monitored data. The site parameters to be documented will vary by program. The site characteristics description should include:

- General building configuration and envelope characteristics such as building floor area, conditioned floor area, number of building floors, opaque wall area and U-value, window area, and solar heat gain coefficient;
- Building occupant information such as number of occupants, occupancy schedule, and building activities;
- Internal loads such as lighting power density, appliances, plug and process loads;
- Type, quantity and nominal efficiency of heating and cooling systems;
- Important HVAC system control set points;
- Changes in building occupancy or operation during the monitoring period that may affect results; and
- Description of the energy conservation measures at the site and their respective projected savings.

Specify M&V Method: The M&V method chosen for the project should be specified. M&V methods should generally adhere to the applicable IPMVP protocol for the defined level of rigor. Considerable latitude is provided to the evaluators to develop a SSMVP, which may be a combination of the IPMVP options or a verification only approach where applicable, such as the TRM Deemed Measures. Section 4.1.2.2 provides overall guidelines for developing M&V plans by measure/technology.

Specify Data Analysis Procedures and Algorithms: Engineering equations and stipulated values, as applicable, should be identified and referenced within the SSMVP. Engineering calculations should be based on the TRM for the Deemed and Partially Deemed Measures. Equations and documentation supporting baseline assumptions should be provided. This is a key component of an SSMVP, in addition to the application documents. Often, data are collected without a clear understanding of its intent. This can result in either extraneous data collection and/or missing data during the data analysis step. Fully specifying the data analysis procedures will help ensure that an efficient and comprehensive SSMVP is presented.

Specify Field Monitoring Data Points: If any actual field measurements are planned, they should be specified, including the sensor type, location and engineering units.

Specify Verification and Quality Assurance Procedures: Data analysis procedures to identify invalid data and treatment of missing data and/or outliers must be provided. This should include quality assurance procedures to verify data acquisition system accuracy and sensor placement issues.

Specify Recording and Data Exchange Formats*:* Data formats compliant with the data reporting guidelines described in Section 6 should be specified.

4.1.2.5.2 On-Site Sampling of Installations

As specified in the Sampling and Uncertainty Protocol (see Section 4.4), the sample designs are used to determine the statistically significant sample of projects to be inspected, in order to estimate the population parameters with a certain level of confidence and precision. This section presents a few options for determining the number of installations to measure and verify, once on-site. The actual sample of installations will vary from project site to project site; thus, the options presented below represent only a few of the selection criteria possibilities. The on-site sampling protocols, however, should be specified to some extent in the program specific EM&V Plans.

As with any sample design, one option is to take a census of all the installations on-site. This is to be done in cases where only a few or one installation was made at the particular project site, or when the variance is large and impacts are high. For instances where multiple installations were made (i.e., 100 lighting installations in one office building), a sample of the various installations should adequately represent the entire site statistics. Samples of measures selected for monitoring at a particular site shall be representative of all measures at the site and shall be selected at random. Measures within a building are often grouped according to similar usage patterns, thus reducing the expected variability in the measured quantity within each usage group. Within each usage group, the sample unit may be the individual measure, a particular circuit or point of control, as designated by the M&V plan. Systematic sampling with a random starting point is acceptable. The sampling strategy shall address all measures present at the site that are subject to the M&V study. The target relative precision for sampling measures within a site is 90% level of confidence/20% margin of error for measures selected for investigation. The sampling unit (measure, circuit, control point) should be designated in the M&V Plan.

The initial assumption regarding the coefficient of variation for determining the minimum required sample size is 0.5.

The sample, in general, should be representative of the population; this is where stratification will be of great use. Measures with similar operating characteristics and end use patterns should be grouped and the sampling algorithm should be designed to achieve 90/20 confidence/precision for each usage group. For example, lighting retrofits in common areas should be separated from lighting retrofits in individual suites in an office building or AHU motor retrofits should be grouped separately from chilled water pump replacements for C&I applications.

Since a certain degree of uncertainty is expected with any sampling approach, an error band should be specified within which the claimed installations or savings will be accepted. The SWE recommends using a maximum 5% error band. The error band should be calculated based on the connected kW load. If the evaluation adjusted kW (connected load) for any usage group in the sample is within +/-5% of the claimed kW, the project savings should be accepted at the claimed value, else, the calculations should be revised and recalculated by the EDC evaluators.

4.1.2.6 Realization Rates for Verified Gross Savings

Program and portfolio level realization rates are used to measure the percentage of the savings verified relative to the savings reported or expected. The realization rate reflects the actual installation and proper operation of measures reportedly implemented. There will be two levels (EDC and SWE) to this process of determining realization rates. The EDC evaluators, in collaboration with the SWE, will perform the initial evaluation of performance and will gather the data needed to determine how much of the savings reported was actually installed and is performing as reported. Typically this will entail sampling of a population of program participants, field inspection of the sample installations and a review of the installation data files and engineering calculations (at the specified level of rigor) used to report the savings at the individual participant sample[39]. Sampling expectations are described in Section 4.4. The attainment in this sample is the realization rate of the entire population.

The SWE will have the following roles in the process:

- Verify the sampling process of program participants assuring proper segmentation, by program or measure type, and sampling protocols, based on population size and desired confidence and precision, are used to attain a valid sample.
- At the SWE's discretion, specify for EDC evaluators size weighted sampling protocols and rigor levels based on the sampled participant's impact.
- At the SWE's discretion, review complete data files and conduct site inspections of any participants with the EDC evaluator without bias as to the selection or timing of the site visits. It will be important that EDC evaluators disclose evaluation site visits and provide the SWE a reasonable opportunity to review the files and attend the evaluation site visit.

[39] Note: Site visits and installation inspections are expected where applicable.

- At the SWE's discretion, perform statistical sampling of programs or program subsets as may be required to provide assurances that the savings, as reported, is as stated at either the program level or portfolio level.

4.1.3 Calculating Verified Net Savings

The SWE understands that gross savings will be used to determine compliance with statutory targets. Net savings is considered valuable to inform program design and program planning, and potentially future iterations of Act 129 and is to be reported, but will not affect compliance.

The Commission directed that EDCs "shall go forward without a Net-to-Gross (NTG) ratio (and adjustment) for the first year."[40] Beyond the first year, the Commission stated that "NTG research is appropriate for directing Act 129 program design and implementation" and "directs the EDCs to collect data necessary to determine the NTG ratio for their programs and to apply the ratio when determining the cost-effectiveness of future modifications of existing programs. For those EDCs that have collected sufficient NTG data to calculate NTG ratios, as determined by the SWE, they are to use those ratios immediately to calculate the TRC for future Act 129 program modifications or changes. For those EDCs that have not collected NTG data to date, they shall… begin collecting NTG data and apply NTG ratios… for Act 129 program modifications or changes." In regards to compliance, the Commission stated that " "NTG ratios will not be used to determine whether the EDCs met their energy and demand reduction targets for the June 1, 2009 to May 31, 2013 program period"[41]

NTG estimates are critical for DSM program impact evaluations. They allow the EDCs to determine the portion of gross energy savings influenced by and attributable to their DSM programs, free from the result of other influences. These savings are called net savings because they are net of the savings that would have occurred in the absence of the program. There are two primary factors that differentiate net savings and NTG ratios from gross savings and realization rates. These factors are free-ridership and spillover. Calculating net savings is more an art than science. Essentially, the goal is to separate out the influence of particular energy efficiency programs from other influences that determine participant and non-participant behavior and decisions. NTG ratios establish the proportion of savings directly attributable to a specific program. The NTG ratio scales the estimated program savings to reflect savings only from measures that were installed because of the program. This effectively removes free-riders, participants who would have installed the measure in the absence of the program, and adds program spillover. Spillover refers to, participants who installed additional measures because of the program's influence but did not seek any of the program incentives. The NTG is defined as:

$$NTG = 1 - FR + SO$$ **Equation 4-1**

[40] 2009 TRC Order, page 25.
[41] 2011 TRC Final Order, page 23-26.

Where,

FR = Free-ridership quantifies the percentage of participants who would have implemented the measure in absence of the EDC program

SO = Spillover quantifies the percentage reduction in energy consumption or demand (that is, additional savings) caused by the presence of the EDC program, but which the program, affects by customers investing in additional energy-efficient measures or activities without receiving and financial incentive from the program.

In order to isolate the effect of a program to be evaluated (to attribute causality), the influence of various factors must be considered. Only the impacts caused by the program should be included in the net savings estimates. There are three major issues that need to be addressed in order to evaluate the effect of a program: (1) the baseline energy use that would have occurred but for the existence of the program, (2) free-ridership, or program participants who would have implemented the efficiency measure in the absence of the program and (3) spillover resulting in additional program savings that would not have occurred without the program's influence.

4.1.3.1 Free-ridership and Spillover

Free-riders are program participants who would have implemented a program measure in the absence of the program. The program can also affect when a participant implements an efficiency measure, the level of efficiency of the equipment installed, and the number of units installed. Developing a good estimate of free-ridership can reduce the degree of uncertainty associated with net energy and demand savings. Spillover is the amount of additional savings from customers investing in additional energy-efficient measures or activities due to their program participation or influence. No program costs are associated with spillover savings, but there are energy saving benefits, increasing program benefit-cost ratios.

The net impact from an efficiency program is determined by adjusting the gross savings for those savings that would have occurred in the absence of the program (free-riders) and those savings that resulted from the program but were not counted in the claimed results (spillover). Program evaluators can use several methods to assess the NTG ratio including: self-report surveys, econometric methods, and market share approaches. The SWE Team recommends that a survey-based approach be implemented for determining free-ridership and spillover rates.

Survey-based stated intentions are one method to estimate free-ridership, by asking participants a series of questions on what they would have done in the absence of the program. The SWE Team will assist the EDCs in developing attribution surveys for net impact analysis that allows for the assessment of free-ridership, participant spillover and non-participant spillover rates. The SWE Team recognizes that great care must be taken when developing the battery of questions used to measure free-ridership and spillover, and the successful research approaches used in such states as Massachusetts and

Wisconsin can serve as templates for the Pennsylvania approach. The SWE Team will also provide a list of guidelines for data that should be collected for free-ridership analyses..[42]

The SWE Team recommends that standardized sampling techniques, data collection approaches, survey questions, survey instruments, and analysis methodology be used by all EDCs. This standardization is important for achieving consistency in the methodology used to estimate free-ridership and spillover impact factors. An example of a standardized approach to determining free-ridership rates is included in Appendix B.

The responses to a sequence of free-ridership questions will be used collectively to compute an overall free-ridership score for each measure. It is very important that more than one question be used to determine the level of free-ridership. A program participant is typically considered a full free-rider when the respondent:

- Had already planned or ordered the measure prior to program participation;
- Had the extra money available in the absence of program funding, and would have paid for the entire measure cost;
- Would have specified the same quantity of measures; or
- Would have specified the same efficiency of measures.

If the respondent would have specified more than 0 percent and less than 100 percent of the measures, they should be considered partial free-riders. Free-ridership estimates are typically developed such that the statistical precision at the measure category level (lighting, HVAC, motors, etc.) is 90 percent confidence with a 20 percent precision range and at the program level is 90 percent confidence ±10 percent in precision.

A measure within a participating project will be considered part of the participant spillover adoption if (1) the measure is program-qualifying, or above code; (2) the project did not receive any rebate for the measure (through any channel); and (3) the respondent reports a sufficient level of program influence to indicate they would not have installed the measure (or the same number of measures) in the absence of the program.

Examples of spillover include:

- Program participants adopting additional measures without an incentive.
- Consumers acting on the program influencing available energy-using equipment in the marketplace.
- Changes brought about by the more efficient practices of architects and engineers due to the program, ultimately forcing consumer behavior into desired patterns.
- Changes in nonparticipants behaviors resulting from program direct marketing or changes in stocking practices.

[42] Guidelines will be developed and discussed at future TWG meetings.

Trade ally surveys and reports can also help to determine spillover. The following responses on trade ally reports should be considered when determining the non-participant portion of a program's spillover impact:

- Installed measures in non-participating projects that were for the same measures funded in participating projects, or above code;
- Installed these measures without a program incentive; or
- The program exerted a sufficient degree of influence on this decision to indicate they would not have installed the measures in the absence of program participation.

4.2 With standardized methods to estimate free-ridership and spillover, the EDCs will also be able to calculate net program savings in a consistent and cost-effective manner. The use of a standardized methodology will ensure that the SWE can conduct comparisons of program successes (i.e., net savings) achieved across all EDCs and the state as a whole. Guidance on Process Evaluation

The primary purpose of the process evaluation is to provide an assessment of one or more program-related characteristics in order to provide specific and highly detailed recommendations for program changes. Typically, recommendations are designed to affect one or more areas of the program's operational practices, such as marketing, program delivery bottlenecks, internal communications or the incentive application process. Process evaluations are a significant undertaking, designed to produce more efficient and more cost-effective programs.

The process evaluation is an important tool in the evaluation toolbox. The process evaluation consists of in-depth examinations of the design, delivery, and operations of energy efficiency programs in order to improve the ability of the program to achieve energy savings and accomplish other program goals. The process evaluation also provides a vehicle for sharing program design and operational improvements with other professionals in the field. When process evaluation results are shared with other energy efficiency professionals, these professionals can assess the relevance of the evaluation findings and recommendations to their policies, programs, and program portfolios. This is especially true for program designers and managers who may want to determine if the evaluation results can be used to improve the cost-effectiveness and operational efficiency of their programs. Listed below are examples of how the results of process evaluations can be used by decision-makers:

- Improve program performance with respect to internal administration and communications, promotional practices, program delivery, incentive levels, and data management;
- Provide a means of improving customer satisfaction and identifying market threats and opportunities;
- Provide information to regulators and other interested parties that [energy programs] are being implemented effectively and modified or refined as necessary; and
- Provide a means of contributing to industry-wide knowledge so that other EDCs can improve their programs.

This Process Evaluation Protocol is designed to provide the PA PUC and other stakeholders a level of assurance that there is a minimum set of standards for process evaluations across the EDC portfolios and allowing the necessary flexibility and control for program administration and process evaluation management.

The process evaluation uses interview and survey techniques to describe and assess program operations, which can be compared to original design intent, as well as to measure participant satisfaction and program performance, which can be analyzed to identify gaps between program goals and results. The outcome of this analysis will provide conclusions and recommendations for enhanced program performance. The results of the process evaluations will help:

- To highlight areas for improvement; and
- To identify best practices that that can be implemented on a going-forward basis.

Each process evaluation should have a program-specific or program-group-specific detailed process evaluation plan to guide the evaluation efforts. These detailed plans should include the process evaluation approach, identification of program-specific or program group-specific focus of the evaluation efforts, detailed researchable issues to be addressed, activity timing issues and the resources to be used.

Process evaluations can be conducted at any time within the program design and implementation process. However, there are some key considerations for the timing of these evaluations associated with the Audit Plan.

1) Design Feedback. In many cases, process evaluation can help programs be more effective or have more efficient operational systems before they are placed in the market. When process evaluation experts work with program designers and managers during the early development and implementation period, the evaluation staff can help identify potential problems associated with early program designs or operational practices.

2) Early Implementation Feedback. Early program evaluation feedback to the program designers and managers is an important component of any evaluation. If the process evaluation is conducted early, the evaluation can help the program designers and managers identify and agree upon an early feedback system that allows managers to be aware of early evaluation findings, and take corrective actions where appropriate, before the evaluation report is finalized. These early feedback systems provide the program designer and manager with a method of ongoing interaction with the evaluation professional, to make sure evaluation results are communicated to program management to allow for program improvements as early as possible.

3) Scheduling the Evaluation Efforts. Process evaluations should be conducted for all significant energy efficiency programs early in the implementation process, so that improvements to the programs can be identified and initiated over a period of time in which the results from the

evaluation can influence the program being evaluated. While a process evaluation conducted near the end of a program funding cycle may help future programs (including extensions of current programs), the evaluation should be conducted in time to help the current program more effectively or more efficiently achieve its goals. This means that the process evaluation may need to be conducted after the program's start-up issues have been identified and dealt with by program management, or at which point the program is considered to be in a steady-state mode of operation. Typically, a new program requires about four to six months to move from the early planning and organization phase, into the early market entry phase, and on to a more normal, steady state operational phase. Typically this means that an in-depth process evaluation can be initiated in about the sixth month following program rollout.

The process evaluation effort can examine the effectiveness of the following program components:

- Program design and operational systems,
- Program tracking and information management systems,
- Internal program communications,
- Program delivery organization and staffing,
- Program staff understanding the program's goals and objectives,
- Skill levels needed to implement the program,
- The methods and procedures used to target the outreach efforts,
- The marketing materials and incentive levels used to promote the program,
- Program operational efforts and their relationship to the program theory and logic model,
- The outreach efforts and the structure and content of these efforts, and
- Early program satisfaction and customer service experiences.

These issues can be examined any time during the life cycle of a program. However, some aspects of the process evaluation may be more successfully assessed during the second or third year of a program cycle. For example, evaluations conducted after the first year might focus on:

- Assessing the relationship between the current program services and the needs of the market or participant,
- The program implementation system and its influence on customer perception of the program,
- The influence of the program on customer behavior and actions, and
- Field practices and their effects on energy savings achieved.

Most programs do not need a process evaluation every year of their implementation cycle. As stated earlier, new programs may want to undergo a process evaluation in the first year and involve the program evaluation staff early in the design process. For this reason, programs may want their process evaluation team on board and engaged during the early development efforts, and for conducting the formal process evaluation initiated within the first year.

4.2.1 Research Objectives

The process evaluation's primary objective is to help program designers and managers structure their programs to achieve cost-effective savings while maintaining high levels of market penetration, customer satisfaction and program efficiency and effectiveness. The process evaluation helps accomplish this goal by providing recommendations for changing the program's structure, incentive levels, management, administration, design, delivery, operations or targets. It is expected that process evaluations will be needed both in the early stages of the program design and deployment efforts to provide timely feedback to the EDCs, and over the life of the program as issues are identified.

Because there can be overlap between the information collected in the process, market effects and impact evaluations and associated M&V efforts, the process evaluation efforts should be structured to coordinate with other evaluation efforts to the extent practical.

The process evaluation may take on the challenge of evaluating most, if not all, aspects associated with the design or operations of a program, in order to improve the type and amount of energy resources acquired (directly or indirectly) by that program. The process evaluation plan should address issues applicable to the programs under review, including: program design, program administration, program implementation and delivery, and market response. Each of these issues is addressed in the following sections.

4.2.1.1 Program Design

The following issues pertaining to program design should be addressed in the process evaluation assessment:

- Program design, design characteristics and design process.
- Program mission, vision and goal setting and its process.
- Assessment or development of program and market operations theories and supportive logic models, theory assumptions and key theory relationships - especially their causal relationships.
- Use of new practices or best practices.

4.2.1.2 Program Administration

The following issues pertaining to program administration should be addressed in the process evaluation assessment:

- Program oversight and improvement process.
- Program staffing allocation and requirements.
- Management and staff skill and training needs.
- Program information and information support systems.
- Reporting and the relationship between effective tracking and management, including both operational and financial management.

4.2.1.3 Program Implementation and Delivery

The following issues pertaining to program implementation and delivery should be addressed in the process evaluation assessment:

- Description and assessment of the program implementation and delivery process.
- Clarity and effectiveness of internal staff communications.
- Quality control methods and operational issues.
- Program management and management's operational practices.
- Program delivery systems, components and implementation practices.
- Program targeting, marketing and outreach efforts.
- The level of financial incentives for program participants.
- Program goal attainment and goal-associated implementation processes and results.
- Program timing, timelines and time-sensitive accomplishments.
- Quality control procedures and processes.

4.2.1.4 Market Response

The following issues pertaining to market response should be addressed in the process evaluation assessment:

- Customer interaction and satisfaction (both overall satisfaction and satisfaction with key program components and including satisfaction with key customer-product-provider relationships and support services).
- Customer or participant energy efficiency or load reduction needs and the ability of the program to provide for those needs.
- Market allies interaction and satisfaction.
- Low participation rates or associated energy savings.
- Market allies' needs and the ability of the program to provide for those needs.
- Reasons for overly high free-riders or too low a level of market effects, free-drivers or spillover.
- Intended or unanticipated market effects.

4.2.2 Data Collection and Evaluation Activities

Process evaluation efforts can include a wide range of data collection and assessment efforts, such as:

- Interviews and surveys with an EDC's, designers, managers and implementation staff (including contractors, sub-contractors and field staff).
- Interviews and surveys with trade allies, contractors, suppliers, manufacturers and other market actors and stakeholders.
- Interviews and surveys with participants and non-participants.
- Interviews and surveys with technology users.
- Interviews and surveys with key policy makers and public goods charge stakeholders.
- Observations of operations and field efforts, including field tests and investigative efforts.
- Operational observations and field-testing, including process related measurement and verification efforts.
- Workflow, production and productivity measurements.
- Reviews, assessments and testing of records, databases, program-related materials and tools

used.

- Collection and analysis of relevant data or databases from third-party sources (e.g., equipment vendors, trade allies and stakeholders and market data suppliers).
- Focus groups with participants, non-participants, trade allies and other key market actors associated with the program or the market in which the program operates.

4.2.2.1 Reviews of Management and Program Operations

The process evaluation can conduct a review of program operations and management efforts. For example, this review can include discussions with program staff detailing their efforts, activities and responsibilities, the steps they take to keep the program on track, and discussions with program managers and supervisors to review operational responsibilities, activities, and monitoring efforts. The process evaluation can then chart the management and operational systems and compare those to the program theory and logic models, to see if program management and operational efforts are consistent with the primary program goals. Such reviews can help the evaluator assess what should be considered to improve the management and operations of the program.

4.2.2.2 Examinations of Customer Handling and Service Delivery Efforts

Customer handling and service delivery are core components of most energy programs. How customers are handled and the way in which services are delivered can directly impact the number of projects implemented and significantly influence customer satisfaction. Often customer handling and service delivery efforts directly affect the ability of the program to draw-in participants who take the actions needed to produce energy saving. A review of these procedures can include discussions with project staff, observation of interaction with the customer or participant, and conducting interviews or surveys with a sample of participants.

4.2.2.3 Interviews with Program Management, Administrators, and Implementers

Program management and staff can be a valuable source of information, as they typically know the program better than anyone. Interviews with lead program planners and managers, their supervisors, and a sample of program staff, including both central staff and field staff, can help the evaluator assess the program design and operations in order to make recommendations for changes that will improve the ability of the program to cost-effectively obtain energy savings.

Subjects important to cover across different personnel include communications within the program, communications with customers, and communications with stakeholders. In addition the interviews can obtain impressions of the program's strengths and weaknesses and perceptions of the program's successes and the quality of work that can be compared and contrasted with those perceptions from stakeholders and participants. These interviews provide an opportunity to gather recommendations for program improvements from the interviewed personnel.

4.2.2.4 Interviews and Discussions with Key Stakeholders and Market Actors

In addition to program staff, many other individuals are involved in a program including policy makers (e.g. PA PUC staff), utility managers, key stakeholders (e.g. trade associations and tenant groups), and a

variety of other market actors such as product manufacturers, distributors, installation contractors, and service personnel. It can often be useful to interview a sample of these key players in order to obtain their insights into what the program is doing well, and what can be improved.

4.2.2.5 Interviews, Surveys and/or Focus Groups with Participants and Non-Participants

One purpose of virtually all process evaluations is to understand the experience of the customers participating in the program, in order to design program improvements resulting from those experiences. Program participants have valuable perspectives on what aspects of the program work well and what aspects represent barriers to participation or satisfaction. Detailed feedback from participants is also important for determining whether the customer's perceptions of specific program attributes and delivery procedures conflict or mesh with program design and management perceptions. Beneficial detailed feedback can include levels of satisfaction with the program and their participation experience and satisfaction with various elements of the program including; satisfaction with the product(s), the organization, scheduling, educational services, quality of work performed, attitude of site staff, responsiveness to questions/concerns, level of savings achieved, etc.

4.3 Guidance on Cost-Effectiveness Analysis

The cost-effectiveness evaluation plan will address the cost-effectiveness of each EDCs portfolio of programs in accordance with (1) the 2008 Act 129 statute and (2), the Total Resource Cost (TRC) Test Order issued in June 2009 and August 2011. As stated in 66Pa. C.S. § 2806.1(m) – the TRC test is a "standard test that is met if, over the effective life of each plan not to exceed 15 years, the net present value of the avoided monetary cost of supplying electricity is greater than the net present value of the monetary cost of energy efficiency conservation measures." In calculating the TRC test, the efforts of both the impact and process evaluations should be leveraged and supplemented with specific EDC data regarding avoided costs and program expenses to estimate the cost-effectiveness of the programs offered and the portfolios as a whole. The SWE Team will review the calculations and assumptions used by each EDC to verify compliance and to determine the benefit-cost ratios as formulated in the TRC order. Additionally, as the SWE, we will compare these cost-effectiveness test results to other best practices programs and portfolios across the nation in order to benchmark the success of the EDC EE&C portfolios in PA.

4.3.1 Total Resource Cost Test

The Pennsylvania PUC has adopted the *California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects* TRC test definition, formula and components with a few slight modifications. According to the California manual:

> The Total Resource Cost Test measures the net costs of a demand-side management program as a resource option based on the total costs of the program, including both the participants' and the utility's costs.
>
> The test is applicable to conservation, load management, and fuel substitution programs. For fuel substitution programs, the test measures the net effect of the

impacts from the fuel not chosen versus the impacts from the fuel that is chosen as a result of the program. TRC test results for fuel substitution programs should be viewed as a measure of the economic efficiency implications of the total energy supply system (gas and electric).

Benefits and Costs: This test represents the combination of the effects of a program on both the customers participating and those not participating in a program.[43]

The TRC test, as it pertains to PA and the EDC EE&C Plan cost-effectiveness evaluation, are outlined in the following sections.

4.3.2 Pennsylvania Total Resource Cost Test Order

According to the 2009 and 2011 TRC Test Orders the formulae for the net present value (NPV_{TRC}), the benefit-cost ratio (BCP_{TRC}), and the levelized costs are:

Equation 4-2

$$NPV_{TRC} = B_{TRC} - C_{TRC}$$

Equation 4-3

$$BCR_{TRC} = \frac{B_{TRC}}{C_{TRC}}$$

Equation 4-4

$$LC_{TRC} = \frac{LCRC}{IMP}$$

The variables used in the preceding equations are defined in Table 4-5. A more thorough description of the variables, data requirements, and assumptions are discussed in the following section.

Table 4-5: Glossary of Terms for TRC Test

Variable	Definition
NPV_{TRC}	Net present value of the total costs of the resource
B_{TRC}	Benefits of the program
C_{TRC}	Costs of the program
BCR_{TRC}	Benefit-cost ratio of the total costs of the resource
LC_{TRC}	Levelized cost per unit of the total cost of the resource (cents per kWh for conservation programs; dollars per kWh for load management programs)
LCRC	Total resource costs used for levelizing
IMP	Total discounted load impacts of the program

[43] *California Standard Practice Manual: Economic Analysis of Demand-Side Program and Projects*: October 2001; Pg. 18.

In general, the costs and benefits accounted for in Equation 4-2 through Equation 4-4 are listed below. Each of these particular costs or benefits may be comprised of a series of individual expenses depending on the program. This will be addressed in the following section in addition to the requirements as addressed in the data tracking and reporting requirements.

4.3.3 TRC Costs and Benefits

The expenses that are accounted for in the TRC test include:

- Incremental initial and annual program costs including:
  - Cost of equipment,
  - Operation and maintenance,
  - Installation,
  - Program administration, and
  - Customer dropout and removal of equipment (less salvage value).
- Increased supply costs for any periods when load has been increased; calculated for both fuels for fuel substitution programs.
- For fuel substitution programs, costs include the increased supply costs for the energy-using equipment chosen by the program participant.

The potential benefits to an EDC as accounted for in the TRC test include savings from avoided supply costs including reductions in the following costs for periods when load has been reduced:

- Transmission
- Distribution
- Generation

4.3.3.1 Supply Cost Components

Supply costs include transmission, distribution, and generation costs. This can either be a benefit to the EDC for periods when there is consumption reduction or costs during periods of increased demand. According to the TRC Order issued in June 2009, the forecasted generation, transmission and distribution (GT&D) avoided costs should occur in three intervals of five years each for the stipulated maximum assumed program/measure life of fifteen years. The methods for estimating these future costs are summarized below.

4.3.3.1.1 Generation Costs

- *First Five Years*: The first five years of avoided generation costs will be estimated by the NYMEX PJM futures price by "prompt month," two months prior to the filing date. For 2009, this is the May 28th, 2009 data and for 2010 (i.e., program year June 2010 to May 2011) it's the May 27th data. The SWE Team will try to obtain time and seasonally differentiated avoided cost forecasts whenever such data is available.
- *Second Five Years*: The second five years of avoided generation costs will be estimated according to the NYMEX natural gas futures price where the prompt month, converted to electric energy

prices using the spark price spread methodology, two months prior to the filing date, will be used.

- *Third Five Years*: The third five years of avoided generation costs are to be estimated using Energy Information Association Annual Energy Outlook data. Although the data specified is annual, this does not preclude "shaping" data by month, season, etc.

4.3.3.1.2 Transmission, Distribution and Capacity Costs

- The T&D costs are to include the following:
  - Transmission prices, as set by FERC, for each EDC zone
  - EDC distribution rates
  - Generally accepted ancillary service rates, included to the extent known
  - Escalation factors as determined by the U.S. Bureau of Labor and Statistics (BLS), the Electric Power GTD sector, industry index for Electric Power Generation (NAICS 221110)[44]
  - To the extent that GTD costs are not adjusted for loss, adjustments for T&D line losses can be included[45]
- Capacity costs will be estimated according to the following:
  - PJM Reliability Pricing Model
  - Escalation rates determined by the U.S. BLS, the Electric Power GTD sector, industry index for Electric Power Generation (NAICS 221110)
- Prices should be stated on a dollars per MW-day, relative to on-peak demand savings, basis

4.3.3.1.3 Supply Cost Adjustments

- *End-Use Adjustments*: In cases where it is appropriate and feasible, the end-use load shapes of the particular measure should be used to calculate the avoided and/or additional supply costs. Otherwise, the EDC should use the class average consumption profile.[46]
- *GTD Costs*: GTD costs, not so adjusted, can be adjusted for losses.
- *Locational, Temporal, and Zonal Differences*: Zonal based adjustment should be made to the GT&D and capacity costs according to the *PJM State of the Market* report data "Zonal real-time, simple average LMP (dollars per MWh)." Additionally, the natural gas prices in years six through ten should be adjusted according to the basis differential between the Henry Hub as the source and TETCOM-3 as the destination for utilities west of the Susquehanna and Transco Zone 5 as the destination for utilities east of the Susquehanna.
- *Compliance with AEPS Act and Carbon Issues*: The costs of compliance with the AEPS Act that are known and knowable should be included in the TRC cost accounting. However, carbon

[44] The historical average annual growth rate in the BLS Electric Power GTD sector price index for the period 2003 through the most recently available annual data point will be used as a proxy for the rate of escalation between the end of the 2013 program year and the beginning of the EIA AEO in year 11. As PJM RPM, distribution, transmission, or ancillary service cost data become known, it should be utilized in place of the BLS factor. (2011 TRC Tentative Order).

[45] See TRC Order Docket No. M-2009-2108601 issues June 2009; pg. 16.

[46] This stipulation is found on pg. 17 of the TRC order.

reduction expenses are not currently accounted for in Act 129 EE&C Plans.

- *Discount Factor*: For the first year TRC calculations, the EDC specific post-tax weighted average cost of capital (WACC) should be used as the discount rate.[47]
- *Customer Self*-Generating Credits: In cases where customers are self-generating electricity, the full retail rate should be assumed when calculating avoiding energy and capacity cost for the TRC test.

4.3.3.2 Program Cost Components

Program Costs[48]:

- Equipment
- Operation and Maintenance
- Installation
- Customer dropout and equipment removal
- Administration
- Marketing
- Evaluation, Measurement and Verification

4.3.4 Customer Costs and Benefits

Participant incurred costs should include the following:

- All out-of-pocket expenses incurred as a result of participating in a program including:
  - Cost of any equipment or materials purchased;
  - Sales tax;
  - Installation fees;
  - Ongoing operation and maintenance costs; and
  - Equipment removal costs (less salvage value as determined by the market value of either the equipment or materials if it can be repurposed).

The following potential participant benefit should be accounted for in the TRC test.

- Incentives paid by any federal agency (e.g., federal tax credits, ARRA incentives) as benefits.[49]
- Avoided capital and operating costs of the equipment/appliance not chosen in fuel substitution programs.

[47] The discount rate to be used beyond the first year of TRC testing will be addressed in future working stakeholder working group session according to the TRC Order of June 2009.

[48] According to the TRC Order of June 2009, the "Incentive payments from an EDC to a customer will not … be included in the TRC test because such costs are a cost to the EDC and a benefit to the customer that cancel each other out."

[49] Note: the TRC Order of June 2009 specifically excludes the inclusion of state tax credits or Act 1 incentives from the TRC accounting methods of the Act 129 EE&C programs.

When calculating the equipment costs incurred by the participant, the installation scenario effects the cost structure used to determine the expense. For replace on burnout scenarios, where the base equipment being replaced has reached the end of its useful life, the incremental cost for a new device or measure is the additional cost incurred to purchase an efficient device or measure over and above the cost of the standard (i.e., less efficient) efficiency device or measure. On the other hand, in retrofit or early replacement scenarios, where the base equipment being replaced is still fully functional, the incremental cost of the new device or measure is the whole amount of the new efficient device or measure, which is to include all installation costs. In cases where new equipment is being procured (i.e., not replacing older equipment), the equipment cost is calculated as the incremental cost of the high efficiency equipment over the current market or code standard efficiency equipment.

4.3.4.1 Participant Data Requirements

The items particular to each measure or program that will require tracking in order to properly account for all participant costs and benefits include:

- Full measure cost: for calculated equipment costs
- Full baseline measure cost
- *End-Use Adjustments*: If it is appropriate and feasible, the end-use load shapes of the particular measure should be used to calculate the avoided and/or additional supply costs. Otherwise, the EDC should use the class average consumption profile.[50]

4.3.5 Miscellaneous TRC Variables

In addition to the specific EDC and participant costs and benefits, there are a few additional variables which must be defined for accounting purpose. Some of these include:

- Discount rate,
- Planning horizon in five year intervals,
- Program year, and
- Expected measure life, not to exceed fifteen years.

For more information on these variables and any additional stipulations which may arise in future versions, please refer to the latest version of the TRC Order approved by the PA PUC.

4.4 Sampling and Uncertainty Protocol

The energy and demand savings and other evaluation parameters for a project are developed, in conjunction with the TRM and M&V protocols discussed in earlier sections of this Audit Plan, by surveying all or some of the population of projects within a program. There are three major options for surveying a population of projects within a program; these are outlined in Table 4-6.

[50] This stipulation is found on pg. 17 of the TRC order.

Table 4-6: Survey Options

Option	How Many Are Measured & Resulting Precision of Estimates	Rank Order of Contribution to Defensibility
Census:	Measure entire population. Statistical precision is not applicable because you are counting every outcome and, therefore, have a full rather than partial enumeration	Highest
Sample : *Probability Sample:* Simple random and stratified random	Measure a randomly selected subset of the population. Probability of a unit entering the sample is known. Sampling precision depends on the number of items, e.g., participants measured. The more measured, the better the precision.	Varies
Systematic: Any non-random method of sampling	Measure a non-randomly selected subset of the population. Probability of selection unknown. Statistical precision, not applicable. Carefully selected representative samples are sometimes claimed to have properties "similar to" probability samples.	Lowest

Although a survey of the entire population, known as a "census" approach, would provide all of the data necessary to determine the true impact values of an entire program, the reality is that M&V takes a lot of time and money that would likely be better spent elsewhere. Therefore, when used effectively, sampling can improve the overall quality of an evaluation study. By limiting resource-intensive data collection and analysis to a random sample of all projects, more attention can be devoted to each project surveyed.

For an impact evaluation, the goal of the sampling and research design is to create a sound, defensible, unbiased determination of the actual gross and net savings for the overall program. If the sampled projects that are selected to be surveyed follow an efficient sample design, and if the data collection and site-specific analysis is free of substantial bias, then the statistical analysis can provide an unbiased estimate of each population characteristic of interest and a good measure of the achieved statistical precision. The estimates of population characteristics developed from the sample data will generally be close to the true values that would have resulted if the data collection had been carried out for all projects in the population. Moreover, an error bound can be calculated to assess the statistical precision of the results. However, sampling a population does carry the risk of biased estimates, therefore, it is important to pay particular attention to the following issues that can affect the final reported savings[51]:

- Bias in research design, data collection, or analysis; and
- Bias due to non-response, refusals or substitutions to initial sample design and project selection.

[51] *The California Evaluation Framework*. Prepared for the California Public Utilities Commission and Project Advisory Group by TecMarket Works, June 2004.

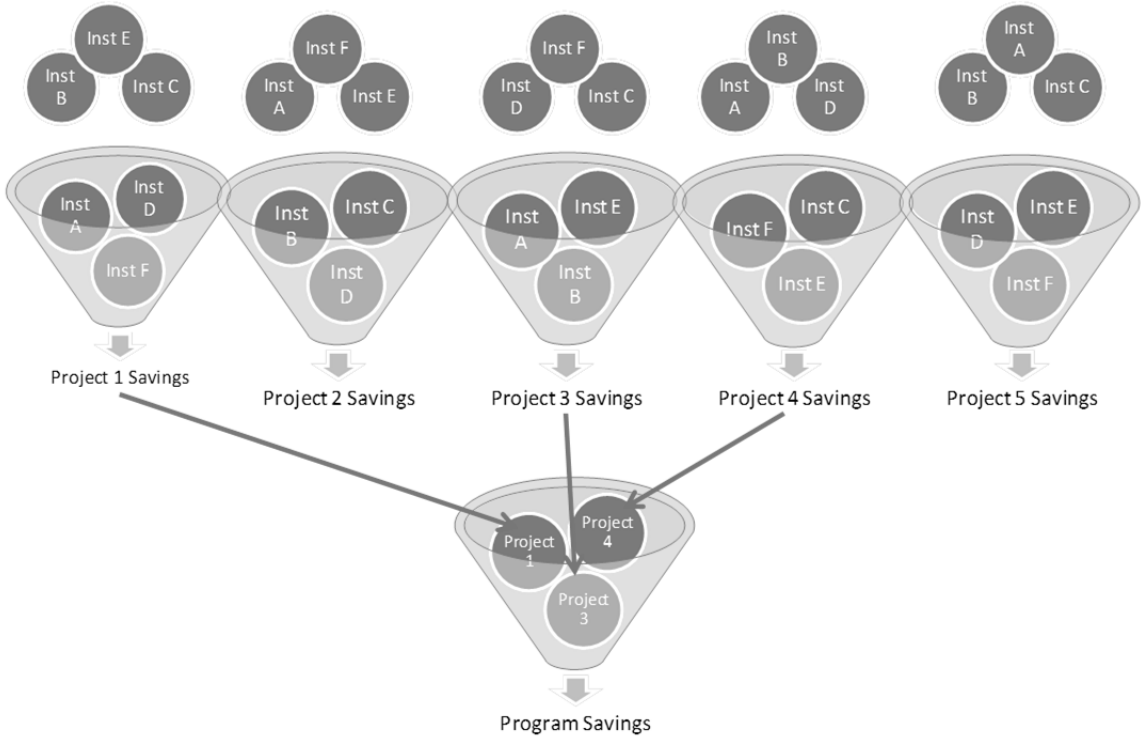
There are several techniques for determining the appropriate sample size of a population to survey in order to ensure a desired level of precision in the results at a specified level of confidence. The next two sections provide an overview of the sampling methods that are appropriate for estimating the savings associated with a defined population. The two methods are: (1) simple random sampling and (2) stratified ratio estimation.[52] Additionally a census can be performed which would survey all projects within a program. This should only be used in instances where the population size is small enough that a census is not cost prohibitive.

To clarify, the goal of the sample design is to determine what project sites are to be visited or surveyed. In other words, out of all of the projects compromising a program, only a few will be selected for site-visits, follow-ups, measurements, etc. The sample design, however, does not specify what to measure or how many measurements to take at one particular site. In general, the variable to be measured on-site is the equipment energy consumption, which is necessary to compute the savings. Thus, the overall variable that is being estimated based on all the M&V activities carried out on the sample is the program savings. In other words, the hypothesis that is being tested by the M&V activities through the sample is the anticipated or tracked gross savings estimates.

Figure 4-3 illustrates the interconnectivity of installation samples ("Inst" in figure), project samples, and population savings. For example, in the figure, out of 5 projects (1-5), only 3 are sampled. And, at each site, only 3 out of the possible 6 installations (A-F) are measured and verified. Yet, the 3 on-site measurements help to estimate the project savings; and the three project savings help to estimate the total impact. Sections 4.4.1 through 4.4.3 outline the framework for sample design and define the protocols for the sample of projects to be selected to estimate the program savings. Site-specific samples will vary from project to project depending on the number of site-specific installations, the variability of savings between installations, and the feasibility of M&V activities.

[52] Note: The sampling protocols developed for the State of Pennsylvania Act 129 Energy Efficiency and Conservation Programs have been adapted from *The California Evaluation Framework* prepared by TecMarket Works for the California Public Utilities Commission and the Project Advisory Group in June 2004.

Figure 4-3: Samples - Project Installations vs. Program Projects



4.4.1 Simple Random Sampling

In simple random sampling, a sample of a given size, denoted *n*, is selected from the projects in the population following any randomized procedure in which all possible subsets of projects are equally likely to be selected. Once the sample is surveyed and the data analyzed, the sample statistics can be extrapolated to represent the corresponding population parameters. The key issue, however, is the performance of these sample statistics as population estimators, which is greatly influenced by sample size and the statistical precision to be expected of the estimator. In evaluation studies, simple random sampling can be effective if the projects in the population do not vary too much in size (i.e., for impact studies, the variability of expected savings is low).

To determine an appropriate sample size for the simple random sampling model, the following variables must be known or assumed:

- *N* = population size,
- *Cv* = coefficient of variation,
- z = z-value, and
- *D* = desired relative precisions.

The coefficient of variation must be estimated by the evaluator based on assumptions concerning the variability of the characteristics being measured in the population. For example, in the case of the impact evaluations, the *Cv* would be based on the variation between projects for expected energy or demand savings. The z-value is based on the desired confidence level assuming a normal distribution. The desired relative precision, *D*, must also be selected by the evaluator. Relative precision is calculated as the error bounds of the sample average over the population mean. Both the desired confidence level and the relative precisions should be based on the M&V protocols established. Given these assumptions, estimates, and values, the following equation can be used to determine the sample size necessary to conduct a simple random sampling of a sufficiently large population.

$$n = \left(\frac{z \cdot Cv}{D} \right)^2$$ **Equation 4-5**

Table 4-7 shows the results of this calculation for various values of *Cv* and *D* given a desired confidence level of 90%.

Table 4-7: Simple Random Sample Size Assuming a Large Population

		Coefficient of Variation, *Cv*						
		0.5	**0.75**	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**
Desired Relative Precision *D*	**0.25**	11	24	43	97	173	271	390
	0.20	17	38	68	152	271	423	609
	0.15	30	68	120	271	481	752	1,082
	0.10	68	152	271	609	1,082	1,691	2,435

However, in cases where there are a small or moderate number of projects in the population, then a finite population correction factor should be applied to Equation 4-5 when determining sample size. The following equation can be used to determine the appropriate sample size for these circumstances.

$$n_f = \frac{n}{1 + \frac{n}{N}}$$ **Equation 4-6**

Here, n is the simple random sample size determined by Equation 4-5 and N is the actual size of the population.

4.4.2 Stratified Ratio Estimation

The stratified ratio estimation sampling technique can be used to address a series of situations for which the simple random sampling method cannot account. Primarily, the stratified ratio estimate sampling methods are better suited for the following situations:

- High variability of annual savings from project to project in the target population; and
- Smaller programs must be consolidated for sampling and survey purposes in order to capitalize on economies of scale for M&V activities.

In cases where the programs must be consolidated for practical M&V purposes, the sample data can be used to provide an unbiased estimate of the average savings per project for each individual program. However, the average savings for the individual programs would generally have poorer statistical precision than the average savings for the portfolio since the individual program results would be based on the sample projects that were drawn from that program.

The stratified ratio estimation method combines a stratified sample design with a ratio estimator. Both stratification and ratio estimation take advantage of information reported in the tracking system and available for each project in the program. The two key parameters in the stratified ratio estimate method are the realization rate and the error ratio. Both variables help to define the relationship between the tracking estimates of savings and the actual project savings. The realization rate is the ratio between the average or total value of the actual savings and the average or total value of the tracking estimates. The error ratio describes the strength of the relationship between the tracking estimates and the actual savings; in other words, it measures the accuracy of the tracking estimates from project to project across the population of projects. The stratified sample design uses the tracked information about the population and projects to add efficiency to the sample design. A stratum is a subset of the projects in the population that are grouped together based on this known information. In other words, a stratification of the population into strata is a classification of all units in the population into mutually exclusive strata that contain projects that share similar characteristics. Under this design, each stratum is sampled according to simple random sampling protocols and the findings can be extrapolated to characterize the entire population.

The statistical issues regarding the stratified ratio estimation method to determine sample size include:

- Estimating the population parameters of interest and calculating the associated confidence intervals,
- Characterizing the population variation when efficiently stratified ratio estimation is to be used,
- Determining the relationship between the expected statistical precision, the population variation and the planned sample size assuming that efficient stratification is used,
- Estimating the required sample size to achieve a desired relative precision,
- Constructing an efficiently stratified sample design, and
- Estimating the relevant population variation from the sample for use in planning future studies.

Appendix C.1 outlines the method required to construct an efficient stratified sampling plan. The primary objectives of the stratified sampling design are to (1) group the projects into several strata based on the tracked savings estimates and (2) specify the number of sample projects to be surveyed from each stratum.

4.4.3 Sampling Roadmap

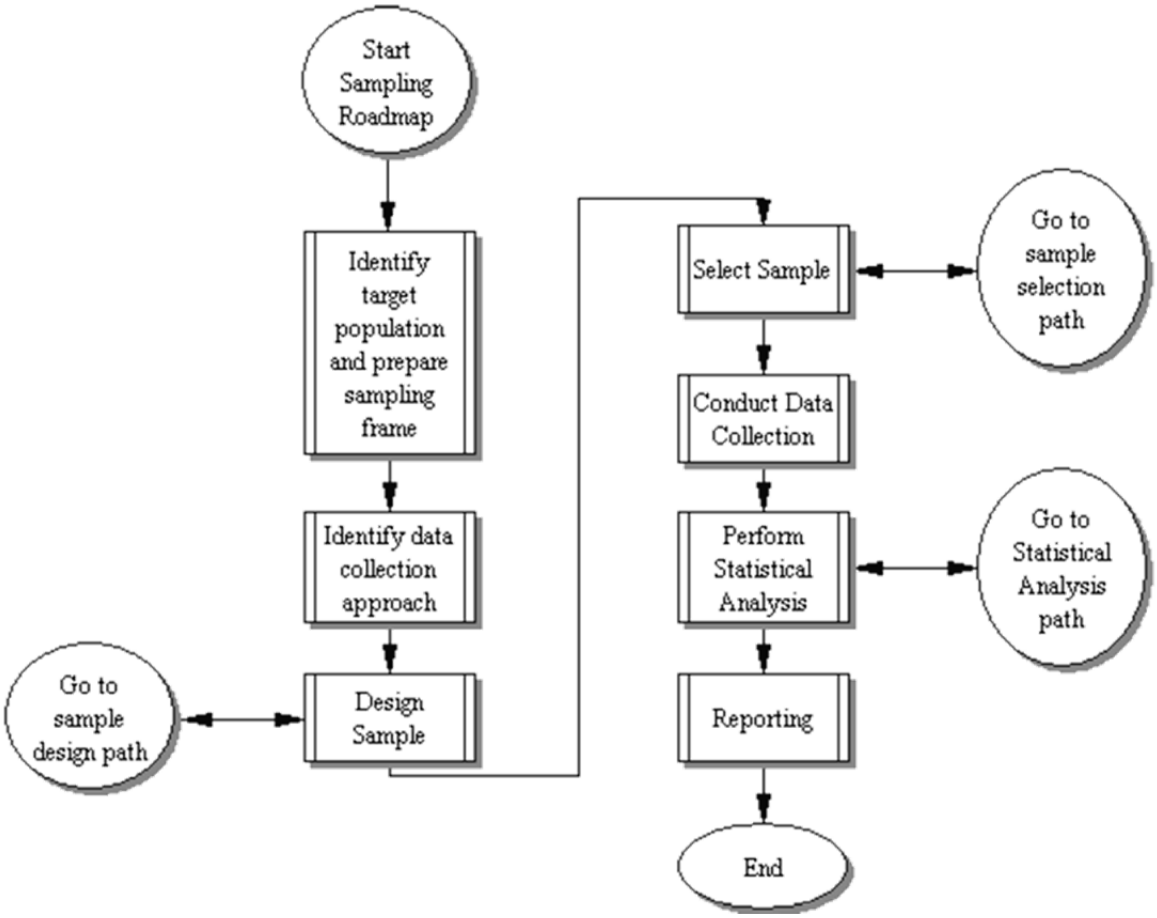
The figure below (Figure 4-4) depicts the typical path of sample design and selection. Each of these stages is described in detail in the following sections by a set of steps and decisions required to create, select and sample a list of projects to evaluate from a program.

Typically, projects are not all installed at the same time and therefore program projects roll-in over the period of program implementation. Therefore, all of the projects are not ready to be randomly sampled until after the completion of the program, which may be too late. An on-going evaluation enables evaluation results to document not only the operations and effects of the program in a timely manner, but also provides feedback for on-going improvement, provides information to support energy efficiency portfolio assessments, and helps improve the cost effectiveness of the evaluation.

To ensure the completion of evaluations within the program cycle and to minimize evaluation costs, a batch approach should be used in the sample design. The batches (i.e., populations) can be customized and sampled every few months, preferably every quarter, during the program implementation year. Each batch (i.e., population) will consist only of the projects installed during the previous batch period. Additionally, in order to ensure that the projects sampled represent the full spectrum of site-specific savings typical of the program, the batches should be stratified based on the estimated site-specific gross savings. The annual sample size for the program should be a projection based on a linear extrapolation of the annual installations in the batch under consideration. The extrapolation should be done using the installation rate of the current and previous batches. Once the extrapolated annual population estimate is calculated, the annual sample size should be calculated for the target precision and confidence, using the equations and steps described in this sampling protocol. The annual sample size should be allocated to the current batch using the relative size of the batch (accumulated number of projects or accumulated savings). At the end of the year, additional samples should be added if the target sample size is not met.

The batch process for sampling is explained in detail in Section 4.4.5. The method prescribed in this protocol is for selecting the number of projects from the population to sample (i.e., the sites requiring M&V activity); it does not address the number of installations to specifically measure once on-site. The on-site sample design is project specific and must be handled on a case-by-case basis.

Figure 4-4: Overall Sampling Roadmap[53]



4.4.3.1 Identify target population and prepare sampling frame

Step 1: Identify the target population and prepare the sampling frame
Prepare database listing all units in the population or program (projects, customers, etc.) with supporting information for each unit.

4.4.3.2 Identify data collection approach

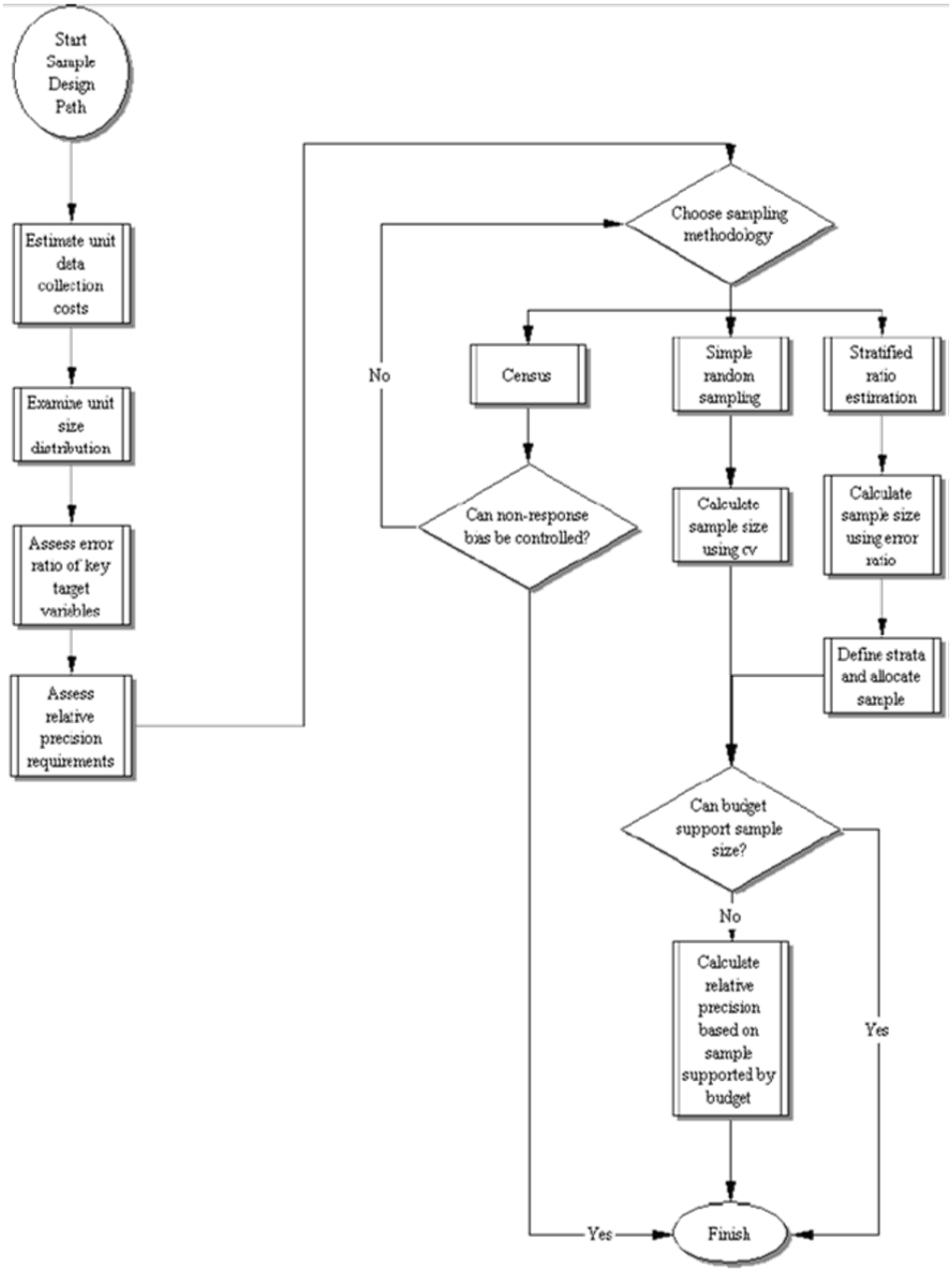
Step 2: Develop the general approach to the data collection and project-specific M&V analysis.
a) Refer to M&V protocols for data collection and analysis approach specific to target population. b) Estimate the average M&V cost for the typical unit in the population. c) If there are less than 50 units in the target population and the data collection costs are moderate, a census may be planned and Steps 3 through 8 may be omitted.

[53] *California Evaluation Framework*: June 2004; Pg. 340.

4.4.3.3 Design Sample

The following figure depicts the stages and decisions of the sample design process.

Figure 4-5: Sample Design Path[54]



[54] *California Evaluation Framework*: June 2004; Pg. 343.

Step 3: Examine the distribution of size of the units in the population.

a) Select a suitable measure of the size of each unit.
   - If the population is a set of projects, the tracking estimate of savings should be used.
   - For non-participant surveys and billing analysis, annual consumption may be used.

b) Calculate the mean, standard deviation and coefficient of variation of size.

c) If the coefficient of variation is less than 1 and the M&V costs are moderate, then a simple random sampling can be used and Step 7 and 8 and Decision 1 can be omitted.

d) This step may be omitted if the stratified ratio estimation is planned[55].

Step 4: Define and assess the error ratio for the key target variable(s) to be collected in the study.

a) Evaluators must start with a default Cv or error ratio of 0.5. The population should be stratified and the Cv can be adjusted for each strata. A higher error ratio is recommended for the large projects strata if the expected variance is high.

b) Cv's and error ratios from similar studies and other jurisdictions may not be used.

c) Cv's and error ratios may be adjusted after substantiated by trended data based on data collected during previous evaluation cycles. Adjusted Cv's and error ratios must be proposed in the evaluator's quarterly or annual sample design with supporting documentation.

d) In most applications, the tracking kWh savings should be used as the x (secondary, known) variable.

Decision 1: Choose the type of sampling methodology, either simple random sampling or stratified ratio estimation.

a) If the error ratio from Step 4 is smaller than the coefficient of variation from Step 3, use stratified ratio estimation unless the reduction in sample size does not offset the added complexity of sample design and analysis.

b) Otherwise, use simple random sampling. If using simple random sampling, complete Steps 5 and 6 and then skip to Step 9.

Step 5: Assess the desired confidence relative precision.

Minimum confidence and precision levels are prescribed in Table 4-8.

Table 4-8: Minimum Confidence and Precision Levels

Segment	Confidence and Precision Level
Residential Portfolio	90/10
Non-Residential Portfolio	90/10
Programs within Each Portfolio	85/15

[55] Note: The SWE recommends a stratified sample design for most programs.

Evaluators may use their professional judgment in the design of the sample as long as minimum requirements are met and are encouraged to exceed minimum requirements when possible. Special consideration should be given to the following situations:

1) Cross-cutting programs that span across both the residential and non-residential sectors must be evaluated as independent programs, one for the residential sector and one for the non-residential sector.
2) The Government, Non-Profit, and Institutional population and the Low-Income population should be evaluated as independent programs if their contribution to their respective sectors (i.e. Residential sector for the Low-Income population, and Non-Residential sector for the Government, Non-Profit, and Institutional population) is over 20%.

Evaluators may propose alternative minimum confidence and precision requirements for programs with special circumstances on an individual basis to the SWE for review and approval. Programs with smaller savings contributions can be sampled at lower precision levels and evaluators should propose the approach with adequate justification in the sample design.

Programs should use stratification to ensure that the sample is efficiently designed and representative of the population by creating homogeneous population groups to the greatest extent possible. Evaluators should use their professional judgment to develop size thresholds for the project strata (e.g. small, medium large). For large project strata, evaluators should ensure that a large portion (approximately 80%) of the savings are evaluated using a high level of rigor. Thresholds are specific to the population studied and can be changed for the sampling batches.

Step 6: Calculate the required sample size to achieve the desired relative precision

a) Refer to Step 5 for expected level of confidence for particular program/population.
b) Calculate corresponding z-value for prescribed level of confidence.
c) If simple random sampling is used, the required sample size is: $n = \left(\frac{z \cdot cv}{D} \right)^2$ provided that the number of units in the population is large compared to the planned sample size. Otherwise, apply the finite population correction: $n_f = \frac{n}{1 + \frac{n}{N}}$.
d) If stratified ratio estimate is used, the preceding equations can be used with the error ratio (*er*) instead of the coefficient of variation (*Cv*).
e) If the indicated sample size is not sufficiently smaller than the population to justify the added complexity of sampling, a census may be used.

Decision 3: If stratified ratio estimation is used, determine whether qualitative strata will be used.[56]

The target population can, if desired, be divided into two or more qualitative strata based on any information that is available for all units in the population (e.g., expected savings, building type, type of measures installed, climate zone, and customer class).

Criteria:

- Stratification by size (savings weights) is strongly recommended. EDC evaluators should use their professional judgment to develop size thresholds for the project strata (e.g. small, medium and large). Thresholds are specific to the population studied and can be changed for individual programs.
- Stratification by measure type or measure group is strongly recommended to the greatest extent possible.
- Innovative stratification methods for sample size allocations are encouraged, e.g. stratification by expected error ratio or expected realization rate range.
- If savings contributions of the Government, Non-Profit, and Institutional and Low-Income populations are under 20% of a program, savings weights may be used to allocate the samples.
- If multiple market segments (e.g. healthcare, retail, schools etc.) are included in the same program, stratification by broad market segments is recommended. EDC evaluators should use their professional judgment to stratify the population. The stratification design should be submitted to the SWE in the annual EM&V plan.
- Qualitative stratification may provide added control over the distribution of the sample
- Qualitative stratification may improve statistical precision if error ratios vary substantially between qualitative strata.
- Sample size in each qualitative stratum should not be too small (e.g., less than 25), so usually only possible if a relatively large sample is planned.
- Qualitative stratification usually adds complexity to the analysis.
- In the analysis, each qualitative stratum can be considered individually and the results combined across the strata.
- Meaningful results can often be developed for some subsets of the sample even if these subsets are not set up as qualitative strata in the sample design.

Step 7: If stratified ratio estimation, develop qualitative strata (optional).[57]

a) Divide the units in the population into unique qualitative strata based on any relevant characteristics (type of building, climate zone, etc.)
b) Tabulate the number of units and total size (e.g., tracking savings) in each qualitative stratum.
c) Assess the error ratio in each stratum (may be equal to the overall error ratio) and allocate the sample to each stratum in proportion to the produce of the total size and the error ratio in each stratum.

[56] This may be appropriate for the government and C&I buildings in participating in joint programs.
[57] This may be appropriate for the government and C&I buildings in participating in joint programs.

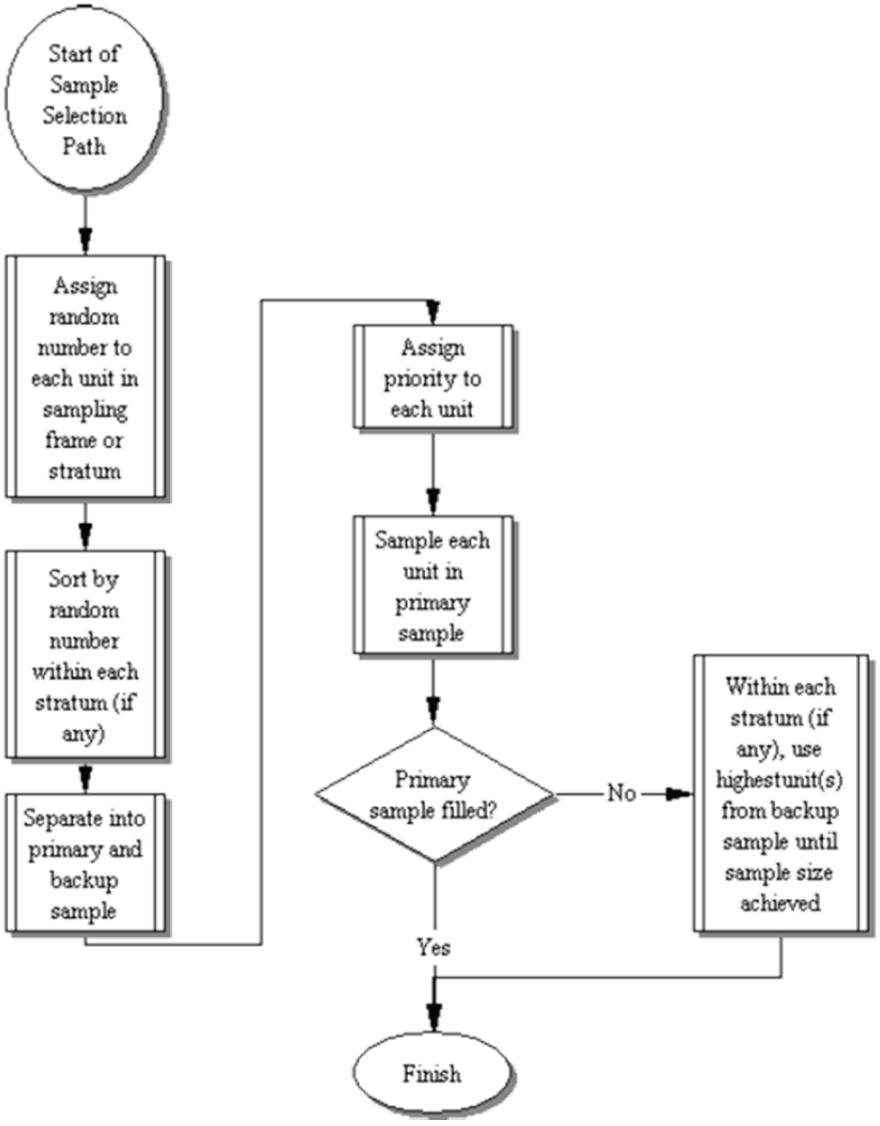
Step 8: Construct the size strata (required under stratified ratio estimation).

1) Choose the number of size strata, denoted *L*, to be used overall or within each qualitative stratum (if used).
2) Use the values of size in the population database to construct the strata overall or within each qualitative stratum (if used).
3) Construct the strata:
   a. Assume a value for the beta parameter of the ratio model, usually β=1 and calculate $\mu_i = \beta x_i$ for each case in the population database. Here x_i is the measure of size of case *i*.
   b. Assume a value for the gamma parameter of the ratio model, usually γ=0.8 and calculate x_i^{γ} for each case in the population database. Here x_i is the measure of size of case *i*.
   c. Calculate σ_0 parameter of the ratio model using the equation $\sigma_0 = er \frac{\sum_{i=1}^{N} \mu_i}{\sum_{i=1}^{N} x_i^{\gamma}}$. Here *er* is the error ratio from Step 4.
   d. Calculate $\sigma_i = \sigma_0 x_i^{\gamma}$ for each case in the population database.
   e. Sort the population database in increasing values of σ_i.
   f. For each case *i* in increasing order, calculate the cumulative sigma $c_i = \sum_{j=1}^{i} \sigma_j$.
   g. Divide the projects in the sorted database so that the sum of the σ_i is approximately equal in each of the *L* strata. In Excel this can be done by calculating $h_i = INT\left(L \frac{c_i}{c_{Ni}} + 0.999 \right)$.
   h. Allocate the overall sample or the sample allocated to each qualitative stratum (if used) equally among the size strata in each stratum.

4.4.3.4 Sample Selection

The following figure depicts the process and decisions required to select the actual projects to be sampled.

Figure 4-6: Sample Selection Path[58]



[58] *California Evaluation Framework*: June 2004; Pg. 349.

Step 9: Assign a random number to each unit in the population database.

Step 10: Sort the units in the sampling farm according to the increasing values of the random number created in Step 9.

Step 11: Assign the priority to each unit in the sampling frame in the order established in Step 9.

Step 12: Designate the units to be included in the primary sample

- Let *n* denote the number of units to be included in the sample. The primary sample is the first *n* units in the sampling frame in the order established in Step 10. In other words, the sample is the units with priority 1 to *n*.
- The replacement or backup units are the next *n* units, to be used in the order established in Step 10. In other words, the units with the priority *n*+1 is the first replacement unit, the unit with priority *n*+2 is the second replacement unit, etc.
- The primary sample should be used initially. A replacement should only be used when a primary sampling unit cannot be used. Replacements should be minimized to control the risk of selection bias. All replacements should be thoroughly documented.
- Evaluator sample sizes must meet sample expectations prior to being augmented by SWE independent site inspections.

4.4.3.5 Data Collection and M&V

Step 13: Data collection and M&V.

For the following components of the M&V process, please refer to the M&V protocols and guidelines outlined in this Audit Plan:

- Recruiting and scheduling
- Using backups
- The sample database
- Quality control

Nested sampling, where on-site inspections are a subset of the sample selected for surveys, is allowed for certain situations:

Table 4-9: Applicability of Nested Sampling

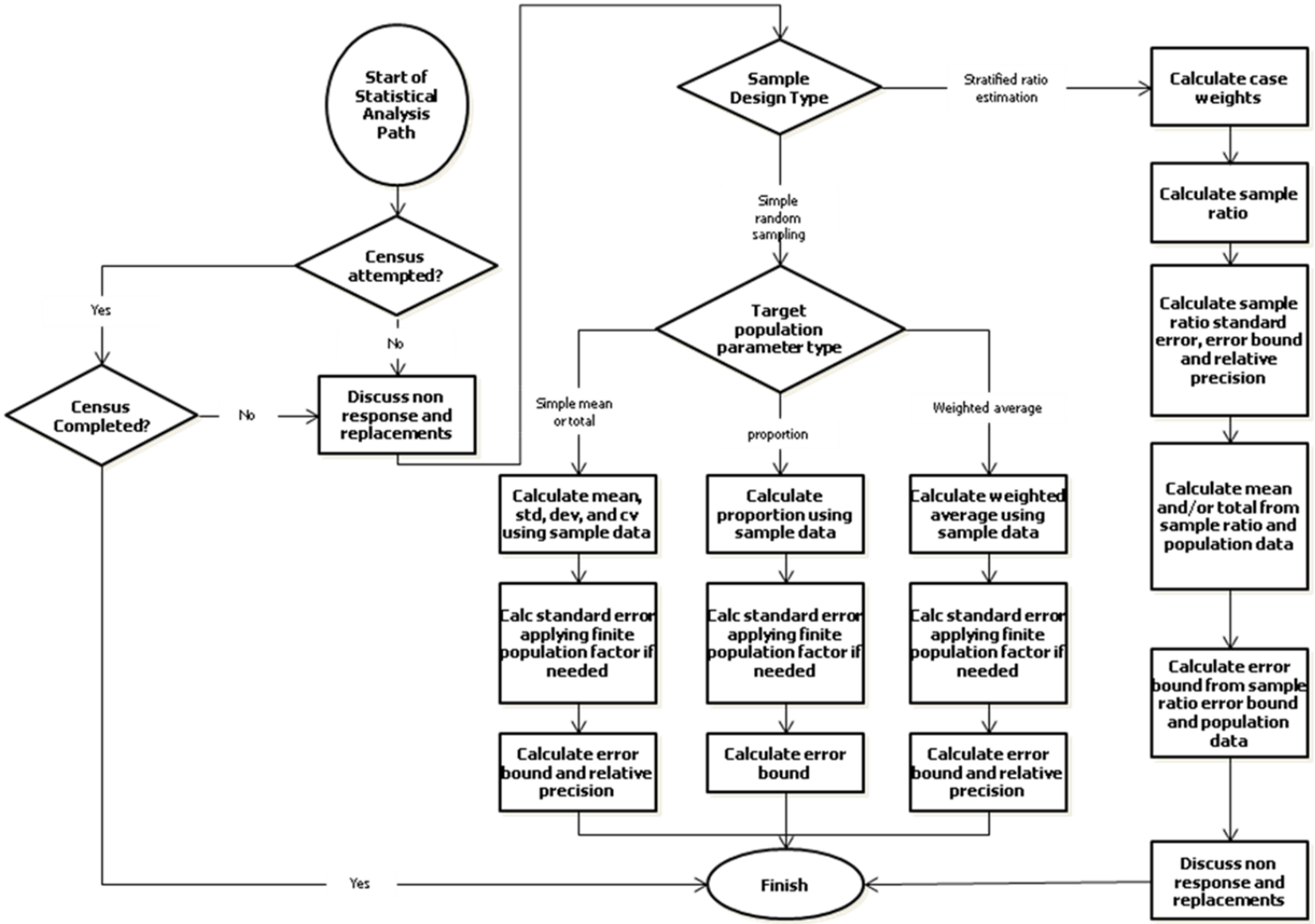
Program Type	Description
Residential Program	Nested Sampling is allowed. The sample size is based on the number of surveys.
Small, Less Complex Non-Residential Program	Nested Sampling can be considered and should be proposed in the EM&V plan for review and approval by the SWE. The sample size is equal to the number of surveys.
Medium & Large, More Complex Non-Residential Program	Nested Sampling is not allowed. The sample size is equal to the number of site inspections.

If nested sampling is used, the sample size of the subset (i.e. the on-site inspection sample) will be based on the evaluator's judgment.

4.4.3.6 Statistical Analysis

The following figure depicts the process of statistical analysis whereby the data collected during the surveys and M&V activities is extrapolated to represent the entire population.

Figure 4-7: Statistical Analysis Path[59]



[59] *California Evaluation Framework*: June 2004; Pg. 353.

Decision 2: Is the target population parameter a simple mean or total, a proportion or a weighted average? Go to Path 1, 2, or 3 respectively.

Path 1: For a simple mean or total.

- Calculate the mean, sample standard deviation and coefficient of variation using the sample data.
- Calculate the standard error using the equation $se = \frac{s}{\sqrt{n}}$. Here *s* is the sample standard deviation and *n* is the number of units in the sample.
- If the sample is more than 5% of the population, multiply the preceding result by the finite population correction factor $\sqrt{1-\frac{n}{N}}$. Here *N* is the number of units in the population.
- Calculate the error bound at the desired level of confidence (according to M&V protocols) using the equation $(z-value)\cdot se$.
- Calculate the relative precision by dividing the error bound by the mean.
- If the population is to be estimated, multiply the sample mean and the error bound by *N*, the number of units in the population.

Path 2: For a proportion.

- Calculate the proportion, denoted $\hat{p}$, using the sample data.
- Calculate the standard error using the equation $se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.
- If the sample is more than 5% of the population, multiply the preceding result by the finite population correction factor $\sqrt{1-\frac{n}{N}}$. Here *N* is the number of units in the population.
- Calculate the error bound at the desired level of confidence (according to M&V protocols) using the equation $(z-value)\cdot se$.

Path 3: For a weighted average or ratio.

- Write the weighted average or ratio in the population as $B = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i}$.
- Calculate the corresponding sample weighted average or ratio using the equation $b = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i}$.
- Calculate the standard error as $se = \frac{\sqrt{\sum_{i=1}^{n} (y_i - bx_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} x_i}$.
- If the sample is more than 5% of the population, multiply the preceding result by the finite population correction factor $\sqrt{1-\frac{n}{N}}$. Here *N* is the number of units in the population.
- Calculate the error bound at the desired level of confidence (according to M&V protocols) using the equation $(z - value) \cdot se$.
- Calculate the relative precision by dividing the error bound by the sample weighted average.

Step 14: Calculate the case weight of each sample project (for stratified ratio estimation only).

- Tabulate the number of population projects in each stratum.
- Tabulate the number of sample projects in each stratum.
- Calculate the case weight as the number of population projects divided by the number of sample projects, and apply the results to each unit in the sample.

Step 15: For any population ratio or proportion, use the sample data and case weights to calculate the sample ratio (for stratified ratio estimation only).

- The population ratio or proportion is defined to be of the form $B = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i}$ where *x* and *y* are any suitable variables.
- The sample ratio is calculated as $b = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i x_i}$. Here w_i is the case weight calculated in Step 14.

Step 16: Calculate the standard error of the sample ratio (for stratified ratio estimation only).

- For each unit *i* in the sample, calculate $e_i = y_i - bx_i$.
- Calculate the standard error of the sample ratio using the equation $se(b) = \frac{\sqrt{\sum_{i=1}^{n} w_i (w - 1) e_i^2}}{\sum_{i=1}^{n} w_i x_i}$.
- Calculate the error bound at the desired level of confidence (according to M&V protocols) using the equation $(z - value) \cdot se$.
- Calculate the relative precision by dividing the error bound by the sample ratio.

Step 17: Calculate the corresponding error ratio for use in future sample designs (for stratified ratio estimation only).

- Assume a value for the gamma parameter of the ratio model, usually $\gamma = 0.8$.
- For each unit *i* in the sample, calculate $e_i = y_i - bx_i$ as in Step 16.
- For each unit *i* in sample, calculate x_i^{γ}.
- Estimate the error ratio between *x* and *y* in the population using the equation $\hat{er} = \frac{\sqrt{\left(\sum_{i=1}^{n} w_i \cdot \frac{e_i^2}{x_i^{\gamma}}\right)\left(\sum_{i=1}^{n} w_i x_i^{\gamma}\right)}}{\sum_{i=1}^{n} w_i y_i}$.

Step 18: If the population mean or total or the primary variable is of interest, and the population mean or total of the secondary variable is known, use the sample ratio to obtain the corresponding estimate and error bound (for stratified ratio estimation only).

- Estimate the population mean or total of *y* by multiplying the population mean or total of *x* by the s ample ratio developed in Step 14.
- Calculate the corresponding error bound by multiplying the population mean or total of *x* by the error bound developed in Step 15.

Step 19: In circumstances where desired, calculate the population ratio for a particular subset of the population (for stratified ratio estimation only).

- Define the indicator variable I_i to be 1 if an element of *S*, and 0 otherwise. I_i
- Define $y_i^* = I_i y_i$ and $x_i^* = I_i x_i$.
- Redefine the population ratio of interest to be $B^* = \frac{\sum_{i=1}^{N} y_i^*}{\sum_{i=1}^{N} y_i}$.
- Estimate *B** as specified in Steps 15 and 6.
- Estimate the population total of *y* for the subset *S* as specified in Step 18.
- Then, if desired, calculate the proportion of the grand total of *y* that is associated with the subset *S*.

If the steps detailed above are followed, then the sample selected for surveying will fit the confidence and precision needs of the specified protocols.

4.4.4 Allocation of Resources to an Evaluation Study

To determine where the greatest benefits lie in conducting impact evaluations on a portfolio of programs, and given the inherent uncertainty and real-life budgetary constraints, there are a few alternative decision processes. The first is based on the Fixed Relative Precision Criterion and the second is the Maximization of Portfolio Savings Precision.[60] Based on the specific EDC program being evaluated and the specified level of confidence and precision, either or none of these methods may be used to help the EDCs allocate M&V funds across all programs. The primary objective of the EM&V process is to meet or exceed the minimum requirements, as specified by the SWE, so that the impacts reported are accurate, reliable and unbiased. So, the particular method of sample selection, fund allocation, etc. is up to the discretion of the EDC so long as the minimum requirements are met and the design methods used are unbiased. This section provides examples of two methods for allocating funds (Fixed Relative Precision Criterion and Maximization of Portfolio Savings Precision). Again, these only serve as examples, and do not preclude the EDCs from using alternate design methods for distributing M&V funds while maximizing the accuracy and reliability of the evaluation findings. Note that these

[60] More information on these two processes can be found in *The California Evaluation Framework* of June 2004.

methods merely estimate the disbursement of funds and the resulting confidence and precision levels that can be expected. For actual sample design and sample selection refer to the previous sections on sampling.

4.4.4.1 Fixed Relative Precision Criterion

For the Fixed Relative Precision Criterion, the following assumptions are made.

1. The objective is to estimate the savings of each program with a fixed relative precision.
2. An efficiently stratified sample of projects will be evaluated from each program and an essentially unbiased estimate of the actual savings of each project will be determined using ratio estimates.
3. The strength of the tracking estimates available in each program is measured by a parameter called the error ratio of the program.
4. Each program has a relative large number of projects.

In this scenario, the sample size that is required for each individual program can be calculated using the simple random sampling method, which is discussed in Section 4.4.1 of this Audit Plan. This method assumes a specified level of confidence and desired precision level for each program being evaluated.

Under the Fixed Relative Precision Criterion method, the confidence and precision of the savings impact are specified for each program in the portfolio. Assuming the program impact estimates are independent of each other, the results of the program specific impacts can be summed to estimate the overall portfolio level impacts. To estimate the impact of the entire portfolio, the following methods are used.

1. Portfolio level impact equals the sum of independent program specific savings.
2. Calculate the error bound for each program by multiplying the desired relative precision by the program savings.
3. Determine the corresponding relative precision according to the following equation: $\sqrt{\sum rp^2}$.
4. Calculate the portfolio level error bound by multiplying the portfolio level savings by the corresponding relative precision (calculated in the previous step).
5. Calculate the portfolio level relative precision by dividing the portfolio error bound by the portfolio savings.

An example of the pooling of program savings to estimate the portfolio level of savings is provided in Table 4-10.

Table 4-10: Pooling the Savings of a Portfolio of Programs

Program	Estimated Savings	Desired Relative Precision	Expected Error Bound
A	50,000	13%	6,264
B	50,000	10%	4,852
C	100,000	6%	6,264
D	500,000	4%	18,792
Total	700,000	3%	21,334

4.4.4.2 Maximization of Portfolio Savings Precision

The other alternative method for managing budgetary constraints and impact evaluation precision is the Maximization of Portfolio Savings Precision. For this method, the following assumptions are made.

1. The objective is to estimate the total evaluation savings of the entire portfolio with the smallest possible error bound.
2. A sample of projects will be evaluated from each program and an unbiased estimate of the actual savings of each project will be determined.
3. The cost of evaluating each sample project is about the same across the projects of the portfolio (i.e., similar M&V methods are being used in the evaluation of each program in the portfolio).
4. The strength of the tracking estimates available in each program is measured by a parameter called the error ratio of the program.
5. Each program has a relatively large number of projects.

Under this method, the overall sample size should be allocated to each program in proportion to the product of the expected savings and the error ratio of the program. In the example provided in

Table 4-11, the maximum desirable sample size is 655[61]. The optimal sample size for each program is determined by its allocation factor. The allocator is calculated by multiplying the expected savings of each program by the assumed error ratio of the same program. The allocation factor for each program is then the program specific allocator over the sum of all allocators. For Program A in the example below, the sample size is calculated according to the following equation:

$$\left(\frac{50{,}000}{580{,}000}\right)655 = 56\,.$$

The relative precision and expected error bound for each program and the portfolio, as a whole, can then be calculated.

[61] Note: This is for illustrative purpose only.

Table 4-11: Optimal Allocation for the Overall Precision

Program	Expected Savings	Error Ratio	Allocator	Optimal Sample Size	Expected Relative Precision	Expected Error Bound
A	50,000	1.0	50,000	56	22%	10,946
B	50,000	0.6	30,000	34	17%	8,479
C	100,000	0.5	50,000	56	11%	10,946
D	500,000	0.9	450,000	508	7%	32,837
Total	700,000		580,000	655	5%	37,280

This Maximization of Portfolio Savings Precision assumes that the cost of conducting each survey is approximately the same; if this is not the case, the method can be modified slightly to account for the discrepancies in survey costs. This alternate method takes into account the following four factors:

- The amount of savings expected from each program.
- The uncertainty about the savings.
- The cost of evaluating each sample project in the program.
- The variability of savings in the population of projects in the program, as measured by the coefficient of variation or the error ratio.

An example of a portfolio of programs with varying unit evaluation costs is presented in Table 4-12. The assumed values include:

- The expected savings
- The error ratio
- The fixed evaluation cost
- The unit cost per project

The sample size for each program is still determined by its allocation factor, but in this method, the allocator is calculated differently. The allocator is calculated by multiplying the expected savings of each program, by the assumed error ratio of the same program, and dividing that product by the square root of the unit cost per project. The allocation factor for each program is still then the program specific allocator over the sum of all allocators. For Program A in the example below, the sample size is calculated according to the following equation:

$$\left(\frac{\frac{50{,}000 \cdot 1.0}{\sqrt{\$100}}}{23{,}811} \right) 655 = 138 .$$

The relative precision and expected error bound for each program, and the portfolio as a whole, can then be calculated normally. In this method, the overall sample size can also be adjusted until the total allocation costs are acceptable or the desired portfolio level relative precision is reached.

Table 4-12: Optimal Allocation with Different Unit Costs in Each Program

Program	Expected Savings	Error Ratio	Fixed Evaluation Cost	Unit Cost per Project	Allocator	Optimal Sample Size	Expected Relative Precision	Expected Evaluation Cost
A	50,000	1.0	$30,000	$100	5,000	138	14%	$43,754
B	50,000	0.6	$30,000	$100	3,000	83	11%	$38,252
C	100,000	0.5	$100,000	$1,000	1,581	43	12%	$143,494
D	500,000	0.9	$100,000	$1,000	14,230	391	7%	$491,444
Total	700,000				23,811	655	6%	$716,944

4.4.5 Dynamic Sampling Methodology

The intent of this section is to explain the dynamic sampling methodology used to measure and report impacts of the conservation programs on an on-going basis. This section provides guidance to design a sampling algorithm for accumulating and evaluating project samples in batches, to meet a target annual sample size requirement defined by the sampling and uncertainty protocol.

Impact evaluation performed simultaneously with program implementation necessitates the need for a dynamic sampling strategy. The dynamic, stratified, batch-wise sampling approach is useful to select a statistically valid quantity of projects, to inspect for purposes of evaluation while the program is being implemented.

There are some benefits inherent in performing implementation and evaluation simultaneously. An on-going evaluation will enable evaluation results to document not only the operations and effects of the program in a timely manner, but will also provide feedback for on-going improvement, provide information to support energy efficiency portfolio assessments, and help improve the cost effectiveness of the evaluation. In this way, lessons learned through evaluation can be flowed back to the program implementation quickly and efficiently.

The dynamic sampling method works by sampling in batches. The batches (i.e., project populations) should be sampled every 1 to 3 months during the program year, preferably each quarter. Each batch (i.e., population) should consist only of the projects installed during that particular batch period. Quarterly batches may be combined under some circumstances. For example, if the participation in a program is minimal in a quarter, dynamic sampling may be used to combine quarterly samples. Additionally, in order to ensure that the projects sampled represent the full spectrum of site-specific savings typical of the program, the batches should be stratified based on energy savings, facility type, fuel type, measure etc. The level of statistical rigor at the program level should be based on the confidence and precision parameters specified in the Table 4-8.

The dynamic sampling method is for selecting the number of projects from the population to sample (i.e., the sites requiring M&V activity); it does not address the number of installations to specifically measure once on-site. The on-site sample design is project specific and must be handled on a case-by-case basis.

The first batch should examine the known quantity of projects completed for the particular program period. With the known quantity of completed projects and the number of days elapsed in the first batch, an annual program population size should be linearly extrapolated based on the time length of the batch. The annual sample size for the program can then be calculated using the criteria and methods described in Section 4.4.3.

The sample size for the first batch should be calculated as the annual program sample size, weighted by the ratio of actual accumulated projects to date, to the quantity of extrapolated annual projects for the program. The sample size for each batch sub-stratum (if additional program stratification is used) should be calculated using the respective sub-stratum energy or demand savings weight (kW or kWh).

Projects to be evaluated should be randomly selected from the first batch and its sub-strata. Each project in the sub-strata should be assigned a random number. Next, the assigned random number for each project should be multiplied by the total peak kW or annual kWh savings of the project to calculate a sampling rank. The appropriate number of projects should then be selected for evaluation, beginning with the largest rank (or higher value) first. All unselected projects should be classified as alternates.

Upon completion of the enrollment period of the first batch, the second round of dynamic sampling should be performed. A revised annual population estimate for the program should be estimated by extrapolating the quantity of projects completed in the batch one and batch two time frames, to the entire year. Next, a revised annual sample size should be calculated at the same confidence/precision level.

The sample size for batches one and two combined can then be calculated by weighting the revised annual sample size, by the ratio of actual accumulated projects to date, to the quantity of extrapolated annual projects. The sample size for batch two is calculated as the difference between the sample size for batches one and two combined and the previously calculated sample size for batch one.

This process should be continued in the same fashion at the end of each batch enrollment period until the end of the program year. At the end of the program year the total samples evaluated should be compared with the target annual sample size, calculated using the actual program population size. If the target sample size is not met, additional samples should be added. If the target is met, the resultant coefficient of variation (Cv_{calc}) should be calculated and compared with the coefficient of variation assumed at the beginning of the sampling method ($Cv_{initial}$). If Cv_{calc} is greater than the $Cv_{initial}$, additional samples should be added to ensure $Cv_{calc} < Cv_{initial}$.

A step by step procedure to construct the sampling algorithm is described below:

1. Establish the Batch 1 timeframe - *1 to 3 months is generally recommended*.

2. Obtain a list of projects installed in the Batch 1 timeframe from the program tracking databases.

3. Estimate the annual program population size using a linear forecast function: *Number of projects in annual population = Number of projects in Batch 1 * 12 Months / Months covered by Batch 1*.

4. Calculate the total sample size (annual) using the project forecast for the annual population, based on the pre-selected confidence/precision level and coefficient of variance. It is likely that the sample size will need to be corrected according to the finite population correction factor.

5. Calculate the sample size for the Batch 1 using a linear function: *Batch sample size = Total sample size * Months covered by batch / 12 months*.

   Calculate the energy savings (or demand savings) weights for all strata in the batch (labeled as stratum 1, 2, 3, 4 in this example). The SWE team recommends evaluators consider stratifying the batches into energy savings groups (using a kW or kWh threshold) or by facility type, fuel type, measure type etc. Savings weights should be calculated as a percentage of the total savings in the Batch:

   S*avings weight for Stratum 1 (%) = Savings (kW or kWh) for Stratum 1 / Savings for Batch 1 (kW or kWh) * 100. Calculated similarly for Strata 1-4.*

6. Calculate sample sizes for individual strata using the savings weights (calculated in Step 5):

   Sample size for Stratum 1 = Savings weight for Stratum 1 (%) x Sample size for Batch 1.

   Sample sizes should be rounded to the closest integer.

7. Assign random numbers to all projects in each stratum. Multiply the assigned random number by the savings (kW or kWh) for the project and calculate a sampling rank for each project.

8. Select samples from each stratum, beginning with the largest sampling rank.

9. Once the second batch of projects has been developed, re-calculate the annual population size combining projects from Batch 1 and Batch 2. Note that timeframes for the batches do not need to be identical.

10. Re-calculate the total sample size (annual unadjusted or adjusted as needed) based on the population size calculated in Step 10 (repeat steps 3-4).

11. Calculate the sample size for the first two batches (combined) using a linear function: S*ample size for the first two batches = Total sample size * Months covered by the first two batches / 12 months*.

12. Repeat steps 6 and 7 for Batch 1 and Batch 2 combined.

13. Sample size for Batch 2 = Sample size calculated in Step 12 – Actual Sample size for Batch 1.

14. Repeat steps 8 and 9 to obtain samples.

Continue this process until the end of the fiscal year. While sampling the last batch, use the measured coefficient of variance from the previous batches. If $Cv_{calc} > Cv_{initial,}$ add more projects in final batch.

4.4.6 Uncertainty

One challenge in evaluating energy efficiency programs is the impossibility of direct measurement of the primary end result—energy savings. Energy savings are the reduction from a level of energy use that did not happen. What can be measured is actual energy consumption after, and sometimes before, the energy efficiency actions. Consequently, the difference between (a) actual energy consumption and (b) what energy consumption would have been, had the efficiency measures not been installed, is an estimate of energy (and demand) savings.[62]

Since impact evaluations seek to reliably determine energy and demand savings with reasonable accuracy, the value of the estimates as a basis for decision-making can be called into question if the sources and level of uncertainty of reported savings estimates are not fully understood and described. While additional investment in the estimation process can reduce uncertainty, tradeoffs between evaluation costs and reductions in uncertainty are inevitably required.

For each EDC's impact evaluation to produce reliable and accurate results, the evaluation process should adhere to a specified set of sampling and uncertainty protocols. Sampling and uncertainty protocols are generally dependent on the type of analysis desired, M&V protocols specified, and inherent uncertainties in measure, demand and usage[63]. (Sampling protocols and practices are discussed in Section 4.4.)

4.4.6.1 Evaluation Uncertainties

Uncertainty can be introduced to the evaluation process in one of two ways: (1) as factors affecting actual savings performance and (2) as savings evaluation M&V uncertainties.

Table 4-13 lists some potential uncertainties common to evaluation studies.

[62] 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>

[63] Note: The elements of evaluation uncertainty developed for the State of Pennsylvania Act 129 Energy Efficiency and Conservation Programs have been adapted from *The California Evaluation Framework* prepared by TecMarket Works for the California Public Utilities Commission and the Project Advisory Group in June 2004.

Table 4-13: Potential Uncertainties in Evaluation Studies

Factors Affecting Savings Performance	Savings Evaluation Uncertainties	
• Instrumentation error • Modeling error • Sampling error • Planned / unplanned assumptions	• Weather • Occupancy level/Schedule • Installed equipment intensity • User demand • Equipment deterioration • Ability of the measure to achieve the intended savings	• Operator behavior relative to specifications • Operator behavior relative to non-ECM equipment usage • Measure implementation effectiveness • Equipment life

Uncertainties can also arise as a result of sampling and evaluation biases. This type of bias can occur when the assumptions of a study are not appropriately implemented. Some of the most common sources of potential bias in impact evaluation include:

- *Non-response and other forms of selection bias*. Selection bias can arise if the sampling frame or sampling plan excludes a particular part of the program population. Non-response bias can occur if any of the designated sample projects are dropped or replaced for any reason. In the non-response case, in order to reduce the risk of bias, another project should be randomly selected from the population to replace the non-response project.
- *Measurement bias*. Measurement bias is represented by a consistent difference between the value of a measurement and the underlying property of the measurement. Examples of measurement bias include: poorly worded survey questions, oversimplified engineering models, inaccurate calibration of measurement instruments, etc.
- *Erroneous specification of the statistical model*. An erroneous statistical model, resulting from a violation of assumptions, can misrepresent the findings of the study.
- *Choosing an inappropriate baseline*. It is important that the baseline estimates appropriately reflect the reality of the situation; otherwise the energy and demand savings estimates can be skewed.
- *Misinterpretation of association as casual effects*. Evaluation studies run the risk of this type of bias when regression and other statistical techniques are used to analyze data that have not been produced in a controlled environment. This can occur, for example, in impact evaluation based on billing data.

Uncertainty to a certain extent is an inherent part of any evaluation, as many of the factors listed above are beyond the control of the evaluator. However, it is expected that evaluators will minimize uncertainties and bias implementing all EM&V protocols as outlined in the Audit Plan and detailed EDC specific EM&V plans, and to quantify and document instances of uncertainty or error that are beyond the control of the evaluator.

4.4.6.2 Documentation of Uncertainties

Evaluation processes should include a thorough documentation of methods, assumptions, uncertainties, and error. This section of the Audit Plan outlines some of the types of documentation that are to be expected in the annual evaluation reports.

The following information is typically used to determine the accuracy of a particular evaluation study or M&V activity where sampling methods are used to estimate savings (see 4.4 for a greater discussion regarding these sampling variables):

- Describe method used to determine sample size;
- Population size;
- Sample size;
- Assumed error ratio or coefficient of variance;
- Estimated evaluation costs at unit, program, and portfolio level;
- Desired confidence and precision levels at the program level;
- Expected savings at the program and portfolio levels;
- Estimated savings at the program and portfolio levels; and
- Portfolio level confidence and precision levels.

The following information is typically used to determine the accuracy of a particular evaluation study of M&V activity where regression models are used to estimate savings[64]:

- Describe procedures used for the treatment of outliers, missing data points and weather adjustment;
- Describe what was done to control for selection bias, if suspected;
- Describe what was done to control for the effects of background variables, such as economic and political activity that may account for any increase or decrease in consumption, in addition to the program itself;
- Describe procedures used to screen data for inclusion into the final analysis dataset. Show how many customers, installations or observations were eliminated with each screen. The reviewer should be able to clearly follow the development of the final analysis dataset;
- Regression statistics: For all final models, provide standard regression statistics in a tabular form;
- Specification: Refer to the section(s) of the study that present the initial and final model specifications that were used, the rationale for each, and the documentation for the major alternative models used. In addition, the presentation of the specification should address, at a minimum, the following issues:
    - Describe how the model specification and estimation procedures recognize and address heterogeneity of customers (i.e., cross-sectional variation);

[64] *California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professions*. April 2009; page 173-174.

  - Describe how the model specification and estimation procedures recognize and address changes in factors that affect consumption over time (i.e., time series variation), apart from program effects;
  - Describe how the model specification and estimation procedures recognize and address the fact that participants self-select into that status, and discuss the effects of self-selection on model estimates, whether or not self-selection is treated explicitly;
  - Describe how truncation within the data and regression towards the mean within the participant population (e.g., within low-income populations) is tested for, the results of this test, and how model specification and estimation procedures recognize and address these issues;
  - Discuss the factors, and their associated measures, that are omitted from the analysis, and any tests, reasoning or special circumstances that justify their omission; and
  - Describe how the model specification can be interpreted to yield the measurement of program impacts.
- Error in measuring variables: Describe whether and how this issue was addressed, and what was done to minimize the problem;
- Autocorrelation: Describe any autocorrelation problems and the solutions specifically taken to address the problem. Specific identification and mitigation diagnostics should be presented, including differing treatment for sub-groups, if any;
- Heteroscedasticity: Describe the diagnostics carried out, the solutions attempted and their effects. If left untreated, explain why;
- Collinearity: Describe procedures used to address the problem of collinearity, and the reasons for either not treating it, or treating it to the level that it was;
- Influential data points: Describe the influential data diagnostics that were used, and how the identified outliers were treated;
- Missing data: Describe the methods used for handling missing data during the analysis phase of the study; and
- Precision: Present the methods for the calculation of standard errors for key parameters, such as gross impacts, net impacts, NTGRs, and key process and market effects measurements.

The following information is typically used to determine the accuracy of a particular evaluation study or M&V activity where engineering models are used to estimate savings[65]:

- Describe the primary sources of uncertainty in deemed and measured parameters used in engineering models;
- Describe the construction of the baseline. Include assessment and description of how the selection of baseline affects the development of gross impacts versus net impacts. Baseline definitions shall be consistent with those used in the net analysis;

[65] *California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professions*. April 2009; page 174-175.

- Discuss efforts to guard against measurement error associated with the various M&V data collection efforts;
- Discuss site selection and potential non-response bias, any tests performed to assess potential bias across and within site measurements, and potential effects of any remaining concerns in this area;
- Describe any potential measurement or bias issues associated with the measurement approaches and tools used as they apply to specific program parameters and estimates:
    - Engineering model bias – systematic under- or over-prediction of effects of a measure by an engineering model;
    - Modeler bias – the systematic under- or over-prediction of effects of a measure by a building energy simulation (e.g., DOE-2) modeler. Also includes the random under- or over-prediction of effects of a measure by a building energy simulation (e.g., DOE-2) modeler;
    - Deemed parameter bias – systematic deviation in a deemed parameter used in an engineering model;
- Meter bias – systematic error in meter and/or sensor;
- Sensor placement bias – systematic over- or under-prediction of measured quantity due to sensor placement (could be combined with above); and
    - Non-random selection of equipment and/or circuits to monitor.

4.4.7 Sample Design Review Procedure

The sample expectations set forth in the sampling sections are intended to be flexible enough to accommodate the wide diversity of EDC portfolios. To verify that the guidelines for sample design are being met, evaluators are required to submit quarterly and annual sample designs to the SWE for review. Generally, sample designs should be submitted with the annual EM&V plan updates. The quarterly sampling models should be submitted with the quarterly reports. The quarterly models will aid the SWE audit activities and will be used by the SWE to select projects for audit. Sample design submissions should include target confidence and precision levels, assumed Cv's and error ratios, and stratification criteria. If evaluators propose alternate confidence and precision levels or Cv's and error ratios, supporting documentation is required. The SWE reserves the right to randomly audit the M&V methods used by evaluators to verify savings for their sampled projects.

5 Audit Objectives and Process

This section of the Audit Plan describes the actions and activities conducted by the SWE to audit the implementation and the evaluation of each EDC program. This will not only include oversight of EDC program delivery mechanisms, but also the verification of all results and evaluation processes conducted by each EDC evaluator to evaluate the credibility and accuracy of the published EDC results. The overall SWE audit findings should be used to inform the EDC evaluation teams when conducting the actual program evaluations. The SWE will use the audit activity findings, which will parallel the EDC evaluation activities, to assess the quality and validity of the EDC program evaluations.

5.1 Audit Objectives

The SWE Audit consists of three major components: impact evaluation audit, cost-effectiveness audit, and process evaluation audit. The following sections of the Audit Plan describe the objectives of the Audit for each of the main components.

5.1.1 Impact Evaluation Audit Objectives

The SWE Team has been retained by the PA PUC to conduct a statewide audit and assessment of the realized savings impact of the EDCs' EE&C programs included in their individual approved EE&C Plans. Under this task, the SWE has segmented the impact assessment into several procedural categories including:

- Standardizing evaluation protocols;
- Reviewing EM&V plans for completeness and consistency;
- Auditing measurement and verification activities for quality control, accuracy, and mitigation of uncertainty; and
- Assessing the achievements of each EDC, and the EDCs combined, in order to evaluate progress towards accomplishing Act 129 goals.

The final SWE assessment of the EDC EE&C plans will be based upon the audit of these evaluation components. A brief description of the individual activities, metrics, and recommended outcomes pertaining to each of these procedural segments are discussed below.

The **standardization of evaluation protocols** are outlined in the Audit Plan and were developed collaboratively with the EDCs and PUC and based upon the protocols outlined in the current TRM and TRC Orders. The standardized protocols outlined in the Audit Plan are for the current program year. Updates to the Audit Plan, based on annual TRM updates, future Commission rulings on such issues as net-to-gross ratios, and best-practices identified throughout the program and audit process, will be approved on an as-need basis and implemented prospectively. It is the intent of the SWE to work with the EDCs and within the confines of the current Act 129 Orders. However, the ultimate goal is to continually work towards the development of a set of standardized protocols throughout the audit process, so that EDC program evaluation may be standardized on a going-forward basis.

The **review of EM&V plans** will be an on-going process. It is the intention of the SWE to work with the EDCs and EDC evaluators in the development of their EM&V plans, in order to address potential issues as soon as possible. Part of this review will include a check of compliance with the expectations and guidelines identified in this Audit Plan. It is the intent of the SWE to work in concert with the EDCs and EDC evaluators to realize the common goal of accurately tracking and reporting realized energy and demand savings.

The SWE **audit of EDC EM&V activities** will include any necessary audit activities required to assess the quality control, accuracy, and uncertainty of EDC EM&V activities and evaluations. The SWE may audit both EDC sampled and non-sampled projects within a program. Additionally, the SWE may randomly accompany EDC evaluators in the field for site inspections, as well as conduct independent audits. On-site, the SWE may either observe program-related M&V activities or take independent measurements. Part of the SWE audit may include an audit of the actual EM&V implementation process as it compares to the approved EM&V plans. The audit will also include a review of engineering and statistical calculations and the transference of data into the EDC specific data tracking and reporting systems. Additionally, in projects of a customized nature, or in instances exhibiting the potential for uncertainty in pre-installation variables affecting the post-installation realized savings, the SWE may review project applications and materials (including, but not limited to, engineering studies, contractor reports, etc.) and participate in pre-installation activities on-site, to determine the appropriates of the given application.

In general, the SWE will focus audit efforts on programs with the most anticipated impact. However, the SWE will review and audit components of all programs; the specific audit activities will be determined according to program design, anticipated impacts, EDC reports and the ultimate value of information (VOI).

During this on-going audit of EDC EM&V activities, the SWE will provide continuous feedback in order to address issues in a timely matter. This will help to mitigate the risk of rejecting any EDC evaluation as it is presented in the final annual report. As mentioned earlier, it is the intent of the SWE to work in concert with the EDCs towards the common goal of accurately portraying the realized energy savings resulting from the EE&C programs.

Ultimately, based upon the specific audit activities undertaken and the findings for a particular program, the SWE will **assess EDC reported impacts,** according to progress towards accomplishing Act 129 goals. These findings may be both qualitative and/or quantitative based upon the audit activities and audit level of rigor. The three possible outcomes of the SWE audit include:

- Full approval of EDC annual evaluation report
- Partial approval of EDC annual evaluation report
- Rejection of EDC annual evaluation report

In the event of a partial approval of an EDC annual report the SWE may:

- Recommend certain project impacts be adjusted based on SWE audit findings; and/or
- Recommend an EDC re-evaluate a particular program based upon the findings and amendments identified by the SWE.

In the event that the SWE rejects an EDC evaluation, the SWE has several options. One option is to recommend that the particular EDC re-evaluate all programs based upon the findings and amendments identified by the SWE. Note that the possibility of cost-recovery for expenses incurred by the EDCs in updating evaluation findings or re-evaluating based upon the recommendations of the SWE is at the discretion of the Commission. Additionally, in the event of a SWE rejection of an EDC evaluation, the PUC may fine the EDC for failure to meet the goals of Act 129.

5.1.2 Cost-Effectiveness Audit Objectives

The cost-effectiveness audit objectives include the verification of the accuracy and reliability of program and portfolio costs, economic benefits and savings reported. Additionally, the SWE will audit the EDCs compliance with the TRC Order in their cost-effectiveness calculations and assess if EDC portfolios comply with the requirements of Act 129.

5.1.3 Process Evaluation Audit Objectives

The process audit objectives include the verification of the accuracy and reliability of program process evaluations. This will involve an audit of the EDC process evaluation scope, survey instruments, and survey implementation. Ultimately, the SWE will assess the validity of the process evaluation findings reported by the EDCs and will make process assertions, highlighting both best practices and areas for improvement, regarding program design and delivery.

5.2 Audit Process

The following figure outlines the various stages of the audit process and the corresponding responsibilities of each party.

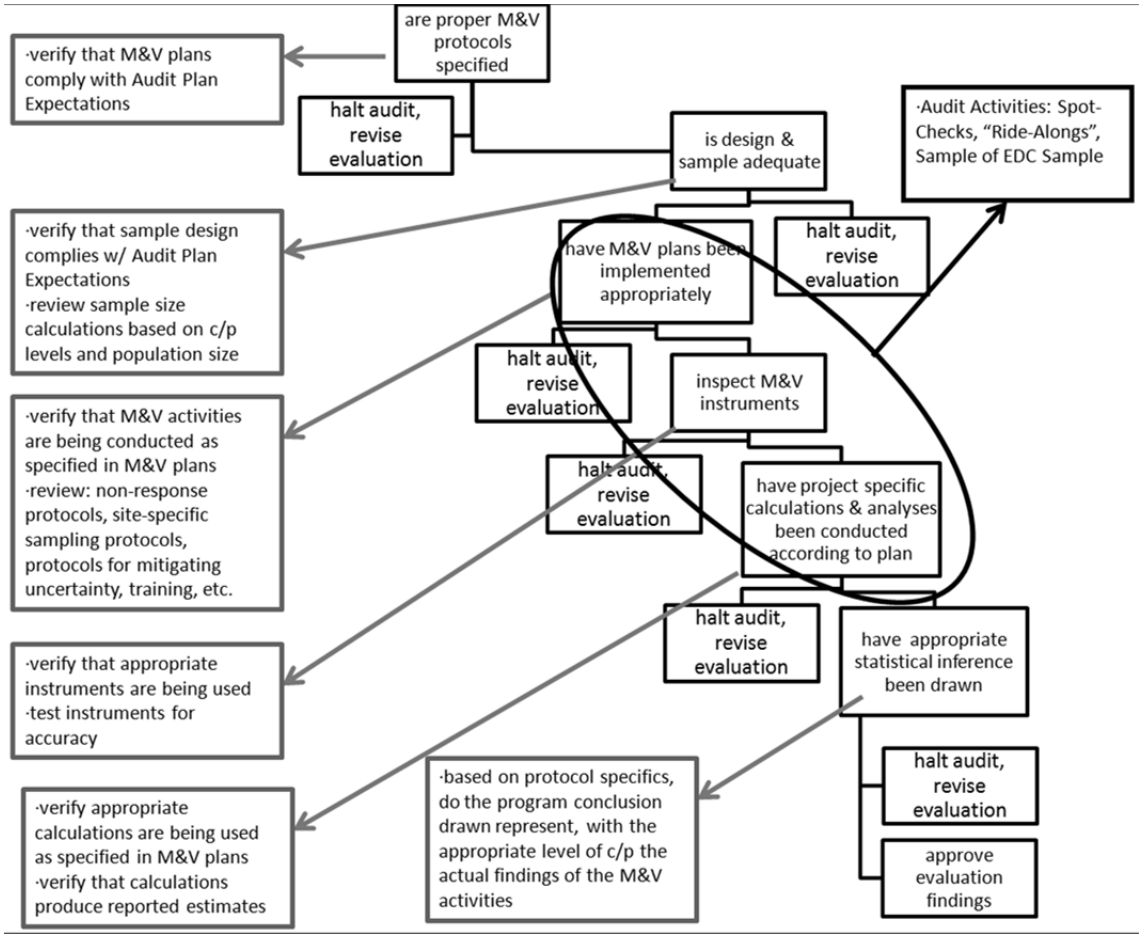
Figure 5-1: Audit Process Flow Chart

Throughout this process, the SWE will refer to the Audit Activities Checklist for the particular program in question, to ensure that all of the SWE responsibilities are upheld. The Audit Activity Checklist includes a list of all the possible activities to be conducted by the SWE Team. It will be customized for each program type so that unique program attributes are properly accounted for throughout the audit process. A template of this Audit Activities Checklist can be found in Appendix D. A more thorough explanation and detailed description of all potential audit activities listed on the checklist are discussed in the following sections of this Audit Plan.

As indicated in Figure 5-1 above, the SWE audit is integrated throughout the EE&C Plan process. This is to help mitigate the risk to the EDCs by allowing the SWE to audit the design and implementation of EM&V activities before or as they occur. This will help the SWE identify potential problems earlier, rather than later, so that the EDCs can effectively incorporate the SWE recommendations so that the likelihood of their ultimate impact reports being accepted is increased. Therefore, the process of evaluation and audit become iterative. This process is depicted in the following figure. Note that the

figure only serves as an example and is not an exhaustive list of the potential activities and findings of the SWE.

Figure 5-2: Iterative Audit Process



5.3 Impact Evaluation Audit

The primary goal of the impact evaluation audit is to certify, with a certain level of confidence and precision, the verified gross energy and demand savings of each EDC. The impact audit will consist of the following:

- Standardizing EM&V protocols;
- Reviewing EM&V plans for completeness and consistency;
- Auditing EM&V activities for quality control, accuracy, and mitigation of uncertainty; and
- Assessing the achievements of each EDC, and the EDCs combined, in order to evaluate progress towards accomplishing Act 129 goals.

The impact assessment task consists of three primary components –preparation, audit and assessment. The preparation component consists of the development of standardized protocols and the review of the EDCs' specific EM&V plans. Section 4.1 of this Audit Plan outlines the SWE's expectations for the

impact evaluations. As part of the audit stage, the SWE Team will monitor and verify data collection, data tracking and reporting quality control. The team will also audit the results of each EDC plan on an on-going basis, with an annual assessment report in accordance with protocols outlined in this Audit Plan.

The SWE responsibilities and activities for auditing impact evaluations as identified in this Audit Plan, and related to monitoring and verifying data collection, quality assurance and results of each EDC plan, include but are not limited to the following:

- Maintaining an evaluation and management database.
- Conducting limited spot field inspections, in coordination with EDCs and Commission Staff, using trained personnel.
- Auditing EDC survey instruments and personnel training.
- Conducting primary data collection or secondary data research to support random spot verifications and savings claims.
- Acquiring data from EDCs and other sources and verifying EDC supplied data.
- Supporting EDC efforts in conducting limited market baseline studies to establish energy efficiency baselines as needed for the impact evaluations of specific programs.
- Analyzing data and reporting findings.

The particular activities, audit sample size, and metrics for assessment will be dependent on the VOI of each specific program. Because of this variability, the Audit Plan lays out the typical or generic audit activities according to the following categories:

- General Audit Activities
- Audit Activities for Deemed Measures
- Audit Activities for Partially Deemed Measures
- Audit Activities for Custom Measures

The specific audit activities falling under these various categories are outlined in the following sections.

5.3.1 General Audit Activities

As stated previously (see Figure 5-2), the SWE Team will first review the EM&V plans of each EDC in order to determine plan adequacy, given the guidelines and expectations as outlined in this Audit Plan. The EDC EM&V plans are very important because they are the foundation for all EDC EE&C evaluations. If this roadmap is not in compliance with the recommendations and expectations of the SWE, then the ultimate evaluation findings of the EDC evaluators may not be verifiable by the SWE. During this review, the SWE Team will check for general compliance issues, which may include:

- Data collection completeness;
- Sample and stratification design;
- Sample selection process;

- Level of rigor for EM&V activities in accordance with the Audit Plan expectations and guidelines;
- Meaningful distribution of M&V funds across M&V activities and programs; and
- Timetable for evaluation activities.

Each of the general audit activity categories will be performed on a sample. However, the size and random nature of the sample will determine the type of conclusions that can be drawn. The three sample options are described below.

1) **Statistically Significant Sample of EDC Sample:** Allows SWE to quantitatively assess the findings of the EDC sample of selected projects (not the program population). Allows SWE to quantitatively assess the validity of particular project findings and qualitatively assess the program evaluation.
2) **Non-statistically Significant Sample of EDC Sample:** Allows SWE to qualitatively assess the findings of the EDC sample as a valid representation of the program population. Allows the SWE to qualitatively assess the validity of particular project findings.
3) **Non-statistically Significant Sample of EDC Program Population:** Would allow SWE to qualitatively assess the finding of the EDC sample as a valid representation of the program population.

Using a combination of management and sampling practices and audit, the SWE Team will verify that the EM&V plans are being followed and that all activities laid out in the plans are being performed. Close coordination will be required between the EDC evaluators and the SWE. This will ensure that the SWE can verify the realization rates calculated by the EDC evaluators. As individual participants are sampled as part of the evaluation's statistical protocol, the SWE will have the option of reviewing files and, after site visits are scheduled, attending the site visit and participating in any measurement and verification activity to assure that the appropriate level of rigor is used. The EDC evaluator will make any adjustments to the savings realization for the participant required by the SWE as part of this process. The SWE may inspect any site desired through this coordinated approach. The intent, however, will be to focus our efforts on high impact projects or measures. The SWE Team plans to allocate a portion of the overall Audit Plan budget to the inspection of such large impact projects. Other participant projects sampled by the EDC evaluator and other data sources used to determine gross savings impacts will be reviewed at the discretion of the SWE. To the extent that inadequate support or evaluation procedures are observed, the SWE may require the EDC evaluator to redo or supplement the evaluation to produce accurate and reliable results.

The SWE Team will evaluate the validity of the utilities reported impacts with the support of the audit findings. The SWE Team will then compile the utility reports in order to assess the overall conservation efforts of the state and the progress made towards the completion of the goals outlined in Act 129.

The following sections contain possible audit activities that the EDCs can expect the SWE Team to perform.

5.3.1.1 Site Inspections

Site inspections are essential to the accurate evaluation of programs and will represent a significant portion of the EDCs' EM&V efforts. Because of the importance of this task, the SWE Team will work closely with the EDCs to ensure that site inspections are carefully planned and executed and that site inspectors have the appropriate experience and training. The SWE Team will audit the following steps in each EDCs site inspection process:

- Training of site inspectors so that they can successfully collect the needed site-specific information.
- Drawing of random sample of sites to be inspected.
- Development of database and site inspection forms.
- Grouping of inspections by geographic location to minimize time allocation, labor and direct costs associated with getting to sites and conducting inspections.
- Contacting sites prior to any visit to ensure availability and so that the resident or facility staff is not 'surprised' by the visit.
- Performing site inspections and entering all needed data into the program evaluation database.

The specific activities included in the SWE's audit will be conducted according to a combination of the following options as detailed in the Audit Activities Checklist. In general, the audit activities fall under two categories. These basic options are described below:

1) **Joint Impact Evaluations:** The SWE may perform "ride-along inspections", in which the SWE accompanies the EDC evaluator on-site inspections to assess performance of the evaluation activities. The EDC evaluator and the SWE will jointly conduct EM&V activities on a sub-set of the EDC evaluation sample with a focus on high impact projects. The site-specific savings may be adjusted based on the findings and recommendations of the SWE.
2) **Independent Evaluation:** The SWE may perform an independent impact evaluation for any sample and related EDC program or sub program impact evaluation that, based on the reasonable expectation of the SWE, has a significant likelihood of inappropriately representing savings. This may include "independent inspections", in which the SWE selects sites in coordination with the EDC evaluator outside of the EDC evaluator sample to verify impacts, while minimizing the bias of influencing the EDC evaluator's results. The SWE may also conduct an independent impact evaluation for any high impact project that the SWE may deem appropriate.

5.3.1.2 Savings Regression Analyses

The SWE Team may audit several of the impact assessments based on regression analysis. Audit activities for these types of evaluations could include:

- Ensuring use of an adequate amount of pre and post-bill summaries;
- Checking the accuracy of data entry;
- Reviewing any statistical analyses performed on the data; and

- Checking for normalization of data for such factors as weather and production rates.

5.3.1.3 Technical Reference Manual and Deemed Savings Formula

As described in Section 3.2.1 of the Audit Plan, the SWE Team plans to chair the TWG set-up, to provide input to the Commission's annual review of the TRM. The team will serve as the technical advisor to ensure that the data used, algorithms implemented, and protocols specified in the TRM are up to date, accurately represent the unique characteristics of the PA EDC service territories and call for protocols that will ensure accurate and reliable assessments.

5.3.1.4 Engineering Models and Simulations

The SWE Team may audit several of the impact assessments based on engineering models and simulations. Audit activities for these types of evaluations could include:

- Reviewing the model or simulation assumptions;
- Verifying that the models and simulations are behaving as intended;
- Checking to ensure that the model inputs accurately reflect the climate and population of the particular circumstance;
- Checking for normalization of data for such factors as weather and production rates; and
- Verifying that the savings estimates produced appropriate reflect reality.

5.3.2 Impact Assessment

Following this impact evaluation process and utilizing the data provided by the EDCs and verified through the audits, the final impact assessment will involve:

- A review and critique of the EDCs' reported energy and demand savings utilizing; field verification, TRM derived savings, approved Custom Measure derived savings, M&V, and large scale billing analyses.
- A review of each EDC's plan to determine whether they are meeting energy savings and load reduction targets.
- Recommendations in the form of best practices, and based on the portfolio of impact assessments for each EDC's EE&C plan, in order to improve realized energy and demand savings.

Ultimately, based upon the specific audit activities undertaken and the findings for a particular program, the SWE will assess EDC reported impacts according to (1) progress towards accomplishing Act 129 goals, (2) cost-effectiveness of portfolios, and (3) effectiveness of program implementation and delivery. These findings may be both qualitative and/or quantitative based upon the audit activities and audit level of rigor. The three possible outcomes of the SWE audit include:

- Full approval of an EDC annual evaluation report
- Partial approval of an EDC annual evaluation report
- Rejection of an EDC annual evaluation report

Possible reasons for rejecting reported impact evaluations are discussed below. Please note that the list is not inclusive and the SWE may identify other reasons for rejecting any portion of the evaluation at any point during the process.

- Failure to correct EM&V Plan issues identified by the SWE. Example of these types of issues include:
  - Inadequate sample design
  - Incomplete data collection activities
  - Inappropriate use of deemed savings or engineering calculations
  - Failure to meet minimum M&V protocol expectations
- Failure to adequately implement M&V Plan. Such issues may include:
  - Inadequate sample drawn
  - Failure to gather all the necessary data on-site
  - Failure to follow protocols specified in approved EM&V Plan
  - Error in calculation or simulation model
- Error in evaluation. This could include:
  - Inappropriate extrapolation of sample findings to estimate population impact
  - Relative precision outside the acceptable bounds
  - High variability between SWE findings and EDC reports

In the event of a partial approval of an EDC annual report the SWE may:

- Recommend that project impacts be adjusted based on SWE audit findings; and/or
- Recommend that an EDC re-evaluate a particular program based upon the findings and amendments identified by the SWE.

In the event that the SWE rejects an EDC evaluation, the SWE may recommend that the particular EDC re-evaluate all programs based upon the findings and amendments identified by the SWE. Note that the possibility of cost-recovery for expenses incurred by the EDCs in updating evaluation findings or re-evaluating based upon the recommendations of the SWE is at the discretion of the Commission. Additionally, in the event of a SWE rejection of an EDC evaluation, the PUC may fine the EDC for failure to meet the goals of Act 129.

5.3.3 Deemed Measures

The impacts of the Deemed Measures are provided in the TRM; the energy and demand savings of these measures are deemed with all energy-related variables stipulated. The only open variable is the number of measure installations. Thus, the audit activities for Deemed Measures will include an audit of the installation verification process. These audit activities will include:

- Compliance with sample design
- Compliance with sample selection
- Compliance with verification protocols, on-site or surveyed

For quality control assurance, the audit activities will include:

- Review of incentive applications;
- Verification of appropriateness of measures given TRM assumptions (i.e., verifying that the measure implemented and project scenario comply with the TRM and program plan); and
- Assessment of the consistency of data between incentive applications and data tracking system.

For verifying the savings, the audit activities will include a review of the formulas used to determine program savings. The SWE Team will verify that the appropriate formulas were used for each particular deemed savings measures.

5.3.4 Partially Deemed Measures

The impacts of the Partially Deemed Measures are estimated using formulas provided in the TRM; these formulas included both stipulated and open variables. The open variables include the number of measure installations and typically a variable designed to represent or describe the end-use loads affected by a particular energy efficiency measure. Thus, the audit activities for Partially Deemed Measures will include an audit of the installation verification process and an audit of the measurement of open variables. These audit activities will include:

- Compliance with sample design
- Compliance with sample selection
- Compliance with verification protocols, on-site or surveyed
- Review of personnel training process
- Compliance with measurement protocols
- Review of measurement instruments
- Quality of data tracking and reporting

For quality control assurance, the audit activities will include a:

- Review of applications;
- Verification of appropriateness of measures given TRM assumptions (i.e., verifying that the measure implemented and project scenario comply with the TRM and program plan); and
- Assessment of the consistency between applications and data tracking system.

For verifying the savings, the audit activities will include a review of the formulas used to determine program savings. The SWE Team will verify that the appropriate formulas were used for each particular partially deemed savings measures.

5.3.5 Custom Measures

The impacts of the custom measures vary from project to project. However, before any project is approved, the savings are typically estimated by the customer, customer representative, or program administrator. Due to the complexity of custom measures and the level of information required to make a reasonable estimate of savings, EDCs may choose to report gross savings based on a variety of

different methods. The EDC evaluator has a heightened responsibility to conduct evaluations such that reported gross savings are verified to an acceptable degree and level of rigor. In some cases, EM&V activities may require the measurement of savings before and after the implementation of the custom measure, while in other cases engineering models and regression analysis is permitted. Therefore, the audit activities for custom measures are typically dependent on the EM&V process selected for the particular category of custom projects.

It is the intention of the SWE to work with the EDCs, EDC program administrators, and EDC evaluators in the development of SSMVPs for custom measures in order to address potential issues before the evaluations are conducted. The SWE reserves the right to review SSMVPs to assess the quality control, accuracy and uncertainty of M&V activities. Guidance on development of SSMVPs can be found in Section 4.1.2.5. For audit purposes, SWE will check if SSMVPs include the following sections:

1. Goals and Objectives
2. Building Characteristics
3. M&V Method
4. Data Analysis Procedures and Algorithms
5. Field Monitoring Data Points
6. Data Product Accuracy
7. Verification and Quality Assurance Procedures
8. Recording and Data Exchange Format

Custom measures still require the verification of installation and will usually involve some sort of open variable measurement, which can be audited. Other audit activities related to custom measures will include:

- Compliance with sample design
- Compliance with sample selection
- Compliance with verification protocols, on-site or surveyed
- Review of personnel training process
- Compliance with measurement protocols
- Review of measurement instruments
- Quality of data tracking and reporting

For quality control assurance, the audit activities for custom measures will generally include a:

- Review of applications and program documents[66];
- Verification of appropriateness of measures; and
- Assessment of the consistency between applications and data tracking system.

[66] Note: The SWE will not be involved in the approval process of individual projects. The review of applications will simply be to ensure that the M&V Plans are implemented appropriately and that the proper data is being collected.

For verifying the savings of Custom Measures, the audit activities will include a review of the formulas or simulations used to determine program savings. The SWE Team will verify that the appropriate assumptions and equations were used for each particular Custom Measure project audited.

The suite of audit activities may differ for each IPMVP option. Specific audit activities for each option are described in Appendix F.

5.4 Cost-Effectiveness Evaluation Audit

Results from the EDCs surveys and M&V activities, evaluation reports, audit, and the statewide impact evaluations will be utilized as inputs to a benefit/cost model and other models, as appropriate, to assess the cost-effectiveness of the EDCs' efforts at the measure, program, sector and portfolio levels where required. In accordance with the requirements of the Commission for determining cost-effectiveness, the EDC's EE&C programs will be evaluated based on the Total Resource Cost (TRC) test. The guidelines for the TRC are stipulated in the 2009 and 2011 TRC Orders. All cost-effectiveness evaluations and assessments will be conducted in accordance with the TRC Order.

5.4.1 Cost-Effectiveness Audit Activities

Possible audit activities pertaining to the cost-effectiveness protocols, calculations and evaluations may include, but are not limited to:

- A review for TRC Order compliance regarding:
  - Formulas
  - Benefits
  - Costs
  - Utility avoided costs assumptions
- A review of EDC accounting practices including:
  - Division of costs and benefits between programs
  - Appreciation/Depreciation rates

5.4.2 Cost-Effectiveness Assessment

The SWE cost-effectiveness assessment will include a review of the benefit/cost (B/C) ratio formulas, benefits, cost and B/C ratio at the EDC project level, EDC plan level, and statewide level. The SWE review of B/C ratios for these various groupings will help to identify the most cost-effective programs across the state. The SWE will also assess these B/C ratios by comparing them to cost-effectiveness calculations from other similar best programs across the country. The SWE will present the findings in a manner that highlights areas of success within the portfolio of EDC projects and that identifies areas of improvement.

5.5 Process Evaluation Audit

As part of the process evaluation and audit, the following thirteen objectives and outcomes reported by the EDCs will be reviewed by the SWE Team:

- Appropriateness of design.
- Appropriateness of participation procedures.
- Appropriateness of application and payment processing activities.
- Accuracy, consistency, and completeness of program records.
- Effectiveness of incentives in motivating action.
- Effectiveness of marketing.
- Effectiveness of internal communications.
- Participant satisfaction with programs.
- Opportunities for process improvement.
- Comparison to best practices.
- Obtain data for assessment of free-riders and free drivers.
- Obtain data for assessment of savings persistence.
- Obtain data for assessment of "spillover" into non-new areas.

5.5.1 Auditable Evaluation Activities

The SWE Team will review the process evaluation plans and activities of each EDC to ensure that they comply with the evaluation framework as outlined in Section 4.2: Guidance on Process Evaluation of this Audit Plan.

The following activities may be audited by the SWE Team to verify the accuracy and confidence of the process evaluation results:

- Review interview and survey materials
- Observe operations and field efforts
- Review workflow, production and productivity measurements
- Review program-related materials and tools
- Review and critique EDC process evaluation reports

5.5.2 Process Evaluation Assessment

The SWE process evaluation assessment will include a review of program design, program delivery, administrative activities, and market response. The SWE review of process findings for these various programs by EDC will help to identify best-practices across the state. The SWE will also assess these process evaluation findings by comparing them to other process and delivery strategies of similar best programs in practice across the country. The SWE will present the findings in a manner that highlights areas of success within the portfolio of EDC projects and that identifies areas of improvement.

6 Data Tracking and Reporting Guidelines

Per the responsibilities outlined in Section 1.2 of this Audit Plan, both the EDCs and the SWE team have reporting requirements that must be met on a quarterly and annual basis. In addition to the reports, the EDCs are responsible for providing the SWE team with the data needed to conduct the impact and process evaluation audits and to populate the public Act 129 website. The reporting and data expectations of the EDCs and the SWE team are outlined in the following sections of this Audit Plan.

6.1 Pennsylvania Act 129 SharePoint Site

The SWE team created a PA Act 129 SharePoint site to improve communication and coordination of activities amongst the SWE team, the BTUS of the PA PUC, the Act 129 EDCs and their respective evaluators, and the Energy Association. This SharePoint Site serves as a depository for documents and data associated with the statewide evaluation of the EE&C Program Portfolios implemented by the seven EDCs. The structure and operation of this SharePoint site complies with the Confidentiality provisions contained in the SWE team contract with the PA PUC and the Energy Association.

The PA Act 129 SharePoint Site contains several different pages. Individual access to each page is based upon assigned administrator privileges and confidentiality of content and the Nondisclosure Agreement signed by all parties and referenced in "Contract Act 129 Statewide Evaluator" (Issuing Office: Pennsylvania Public Utility Commission Bureau of Conservation, Economics, and Energy Planning; RFP-2009-1).

The PA Act 129 SharePoint Site pages include:

- **SWE Site (Home)** which provides a common interface for all parties directly involved in the statewide evaluation efforts and grants access to the Act 129 SharePoint Site. This home page includes the following features: calendar, task lists, technical libraries, report libraries, submission logs, discussion boards, etc.
- **SWE Team Site** where access is restricted to members of the SWE team and the BTUS staff. The purpose of the SWE team site is to facilitate the coordination of activities amongst the various SWE team members, to keep track of assignments, to maintain lists of unresolved issues, etc.
- **Individual EDC Sites** where each EDC has its own password protected site. The EDC specific sites are tailored to fit the needs of the individual EDCs and include such features as submissions library, task lists, memo libraries, etc.

Currently the site is accessible by over 80 unique users who have been approved by the SWE team and the EDCs for which the user works. All requests for new access should be sent to Caroline Guidry (caroline.guidry@gdsassociates.com).

6.2 Public Accessible Website Data Requirements

The SWE team is developing a public accessible website to publish and report on the progress made by the EDCs toward the achievement of their individual Act 129 targets. The website will include basic program information and updates as well as the following:

- Gross reported savings by sector and program
- Verified savings by sector and program
- Total participants by sector
- Measures offered to low-income customers
- Progress towards the achievement of Act 129 targets[67]

Once the website design is approved by the BTUS, the SWE team will provide a data template to each of the EDCs no later than Dec 31, 2011. The data template will be an Excel file with the data fields needed to produce the reports listed above. The SWE team requests that the template provided be used, without modification, to transmit data for the purposes of the public website; the file will be used to automatically update the website according to the data fields specified in the template. This template will be distributed to the EDCs and posted to the PA Act 129 SharePoint site for future access.

6.3 EDC Reporting Requirements

The EDCs are required to submit quarterly and annual reports to the SWE team and the BTUS. In the *Implementation Order* entered January 16, 2009, the Commission noted that Act 129 requires EDCs to submit an annual report documenting the effectiveness of their EE&C plans, the measurement and verification of energy savings, the evaluation of the cost effectiveness of expenditures and any other information required by the Commission.[68] A Secretarial Letter was developed to provide subsequent guidance to the EDCs regarding the Act 129 annual reporting requirements and deadlines[69].

The SWE team has provided the EDCs with a quarterly and annual report template, which is available on the PA Act 129 SharePoint Site. The deadlines for the EDC reports are as follows:

Table 6-1: EDC Reporting Deadlines

Report	Due	Savings Reported
Program Year X, Quarter 1	October 15th	• Quarter 1 Report • Implementation & Evaluation Updates • Gross Savings: Incremental & To-Date • Preliminary Verified Savings
Program Year X, Quarter 2	January 15th	• Quarter 2 Report • Implementation & Evaluation Updates • Gross Savings: Incremental & To-Date

[67] Progress will include figures from program inception to date.
[68] See 66 Pa. C.S. § 2806.1(i)(1).
[69] Secretarial Letter issued on May 25, 2011, at Docket No. M-2008-2069887. Page 1-2.

		• Preliminary Verified Savings
Program Year X, Quarter 3	April 15th	• Quarter 3 Report • Implementation & Evaluation Updates • Gross Savings: Incremental & To-Date • Preliminary Verified Savings
Program Year X	July 15th	• Preliminary Annual Report • Gross Savings as of May 31st • Preliminary Verified Savings
Program Year X	November 15th	• Final Annual Report • Gross Savings as of May 31st • Verified Savings as of May 31st

The preliminary annual reports, the final annual reports, and the quarterly reports shall be filed with the Commission's Secretary and the SWE team via the PA Act 129 SharePoint Site. The Commission will post these reports on its website for public access.

6.4 EDC Data Tracking Requirements

In a Data Request issued October 2010[70], the SWE team requested information and data pertaining to the following:

- General implementation and evaluation information
- Residential Program Data
- Commercial & Industrial Program Data

The SWE team requested that, going forward, the EDCs provide the same type of information with the submission of each quarterly and annual report. All information provided in response to the data request should correspond to activities occurring during the quarter for which the report was submitted. Additionally, the memo included instructions for uploading the data requested to the EDC specific, password protected, PA Act 129 SharePoint Site page.

6.5 Statewide Evaluator Reporting Requirements

The SWE team will submit **quarterly and annual reports** to the Commission with updates on impact evaluations, cost-effectiveness, and process evaluations. These reports will

- Summarize program and portfolio progress to-date for each EDC,
- Summarize energy (MWh) savings and demand (MW) reductions,
- Identify each EDC's savings achievement levels to-date,
- Identify best practices exhibited to date,
- Identify areas for improvements,
- Identify any necessary recommendations for updating targets or expectations based on the current findings, and

[70] The Data Request is available for the EDCs on the PA Act 129 SharePoint Site.

- Provide a summary of audit activities and findings based on the audit work completed during the quarter.

The reports will also include a summary of general activities corresponding to the responsibilities of the SWE team. This could include status of TRM updates, resolutions from Technical Working Group Meetings, summary of guidance memo's issued, etc.

The deadlines for the SWE reports are as follows:

Table 6-2: Evaluation and Audit Deadlines

Report	Due	Savings Reported
DRAFT: Program Year X, Quarter 1	November 15th	• Quarter 1 Report • Summary of EDC progress • Summary of audit activities and findings • Summary of SWE team activities
FINAL: Program Year X, Quarter 1	December 15th	• Final Quarter 1 Report to include comments from BTUS staff and EDCs.
DRAFT: Program Year X, Quarter 2	February 15th	• Quarter 2 Report • Summary of EDC progress • Summary of audit activities and findings • Summary of SWE team activities
FINAL: Program Year X, Quarter 2	March 15th	• Final Quarter 2 Report to include comments from BTUS staff and EDCs.
DRAFT: Program Year X, Quarter 3	May 15th	• Quarter 3 Report • Summary of EDC progress • Summary of audit activities and findings • Summary of SWE team activities
FINAL: Program Year X, Quarter 3	June 15th	• Final Quarter 3 Report to include comments from BTUS staff and EDCs.
DRAFT: Program Year X	January 16th	• Annual Report • Summary of Gross Savings as of May 31st • Summary of Verified Savings as of May 31st • Summary of audit activities and findings • Summary of SWE team activities
FINAL: Program Year X	February 27th	• Final Annual Report to include comments from BTUS staff and EDCs.

6.6 Biannual Workshops

To ensure that all of the EM&V efforts are coordinated and in compliance with the approved Audit Plan, the SWE team will provide workshops twice a year with representatives from each EDC, the Commission Staff, and any other necessary parties as deemed appropriate by the Commission. These workshops will focus on the EM&V "best" practices, evaluation processes, program status and the progress of the audit. The workshops will be designed so that all parties can provide input into the evolution of the audit and evaluation process, as well as discuss identified best practices and areas for improvements. These workshops will help to ensure that all EDCs are performing evaluations of equal quality and substance, as well as implementing programs so as to reach the savings targets in the time allotted.

7 Final Remarks

The primary objective of the EDC energy efficiency programs is to reach the level of savings specified in Act 129 in a meaningful, efficient, and cost-effective manner. It is the desire of the SWE to work closely and collaboratively with the PUC and EDCs in order to develop and implement an evaluation and audit process that will produce significant and standardized impact results, at the lowest cost, so that more funds can be allocated to customer-centric savings activities. As the SWE, however, it is our duty to ensure that the evaluations are accurate and represent the actual impacts of the EE&C program with a certain level of precision and confidence.

This Audit Plan outlines the expected metrics, methodologies and guidelines for measuring performance detailing the processes that should be used to evaluate the programs sponsored by the EDCs throughout the state. It also sets the stage for discussions amongst a TWG of the EDCs, their evaluators, the SWE Team and the PUC. These discussions will help clarify and interpret the TRM, add new prescriptive measures to the TRM and define acceptable measurement protocols for implementing Custom Measures in order to mitigate risks to the EDCs. Our common goal requires kWh and kW savings to be clearly defined, to be auditable and to provide a sound engineering basis for estimating energy savings.

Appendix A. Glossary of Terms

ACCURACY: An indication of how close a value is to the true value of the quantity in question. The term could also be used in reference to a model or a set of measured data, or to describe a measuring instrument's capability.

ACHIEVABLE POTENTIAL: The amount of energy use that efficiency can realistically be expected to displace assuming the most aggressive program scenario possible (e.g., providing end-users with payments for the entire incremental cost of more efficiency equipment). This is often referred to as maximum achievable potential. Achievable potential takes into account real-world barriers to convincing end-users to adopt efficiency measures, the non-measure costs of delivering programs (for administration, marketing, tracking systems, monitoring and evaluation, etc.), and the capability of programs and administrators to ramp up program activity over time.

ADJUSTMENTS: For M&V analyses, factors that modify baseline energy or demand values to account for independent variable values (conditions) in the reporting period.

ADMINISTRATOR: A person, company, partnership, corporation, association or other entity selected by the EDC and any subcontractor that is retained by an aforesaid entity to contract for and administer energy efficiency programs under Act 129.

BASELINE DATA: The measurements and facts describing facility operations and design during the baseline period. This will include energy use or demand and parameters of facility operation that govern energy use or demand.

BASELINE FORECAST: A prediction of future energy needs that does not take into account the likely effects of new efficiency programs that have not yet been started.

BASELINE MODEL: The set of arithmetic factors, equations or data used to describe the relationship between energy use or demand and other baseline data. A model may also be a simulation process involving a specified simulation engine and set of input data.

BASELINE PERIOD: The period of time selected as representative of facility operations before retrofit.

BIAS: The extent to which a measurement or a sampling or analytic method systematically underestimates or overestimates a value.

BILLING DATA: Has multiple meanings. Metered data obtained from the electric or gas meter used to bill the customer for energy used in a particular billing period. Meters used for this purpose typically conform to regulatory standards established for each customer class. Also used to describe the data representing the bills customers receive from the energy provider and also used to describe the customer billing and payment streams associated with customer accounts. This term is used to describe both consumption and demand, and account billing and payment information.

BILLING DEMAND: The demand used to calculate the demand charge cost. This is very often the monthly peak demand of the customer, but it may have a floor of some percentage of the highest monthly peak of the previous several months (a demand "ratchet"). May have other meanings associated with customer account billing practices.

BUILDING ENERGY SIMULATION MODEL: Computer models based on physical engineering principals and/or standards used to estimate energy usage and/or savings. These models do not make use of billing or metered data, but usually incorporate site-specific data on customers and physical systems. Building Simulation Models usually require such site-specific data as square footage, weather, surface orientations, elevations, space volumes, construction materials, equipment use, lighting and building occupancy. Building simulation models can usually account for interactive effects between end-uses (e.g., lighting and HVAC), part-load efficiencies and changes in external and internal heat gains/losses. Examples of building simulation models include ADM2, BLAST and DOE-2.

CAPACITY: The amount of electric power for which a generating unit, generating station or other electrical apparatus is rated either by the user or manufacturer. The term is also used for the total volume of natural gas that can flow through a pipeline over a given amount of time, considering such factors as compression and pipeline size.

COEFFICIENT OF VARIATION: The sample standard deviation divided by the sample mean (Cv = sd/y).

COINCIDENT DEMAND: The metered demand of a device, circuit or building that occurs at the same time as the peak demand of the building or facility or at the same time as some other peak of interest, such as a utility's system load during the average 100 peak summer hours. This should properly be expressed so as to indicate the peak of interest, e.g., "demand coincident with the building peak."

CONFIDENCE: An indication of how close a value is to the true value of the quantity in question. Confidence is the likelihood that the evaluation has captured the true impacts of the program within a certain range of values (i.e., precision).

CONSERVATION: Steps taken to cause less energy to be used than would otherwise be the case. These steps may involve, for example, improved efficiency, avoidance of waste, and reduced consumption. Related activities include, for example, installing equipment (such as a computer to ensure efficient energy use), modifying equipment (such as making a boiler more efficient), adding insulation, and changing behavior patterns.

CONSTRUCT VALIDITY: The extent to which an operating variable/instrument accurately taps an underlying concept/hypothesis, properly measuring an abstract quality or idea.

CONTENT VALIDITY: The extent to which an operating measure taps all the separate sub-concepts of a complicated concept.

CONVERGENT VALIDITY: When two instruments/questions/measurement methods obtain similar results when measuring the same underlying construct with varying questions/approaches.

CORRELATION COEFFICIENT: A measure of the linear association between two variables, calculated as the square root of the R^2 obtained by regressing one variable on the other and signed to indicate whether the relationship is positive or negative.

CORRELATION TABLE (CORRELATION MATRIX): A table or matrix giving the correlation between all pairs of data sets. Row headings are the scores on one variable and column headings are the scores on the second variables and a cell shows how many times the score on that row was associated with the score in that column

COST-EFFECTIVENESS: An indicator of the relative performance or economic attractiveness of any energy efficiency investment or practice when compared to the costs of energy produced and delivered in the absence of such an investment. In the energy efficiency field, the present value of the estimated benefits produced by an energy efficiency program as compared to the estimated total program's costs, from the perspective of either society as a whole or of individual customers, to determine if the proposed investment or measure is desirable from a variety of perspectives, e.g., whether the estimated benefits exceed the estimated costs. See also TOTAL RESOURCE COST TEST.

CUSTOMER: Any person or entity responsible for payment of an electric and/or gas bill to and with an active meter serviced by a utility company.

CUSTOMER INFORMATION: Non-public information and data specific to a utility customer that the utility acquired or developed in the course of its provision of utility services.

CV: See COEFFICIENT OF VARIATION.

DEEMED SAVINGS: An estimate of the reported energy savings or energy-demand savings outcome for a single unit of an installed energy efficiency measure that (a) has been developed from data sources and analytical methods that are widely considered acceptable for the measure and purpose and (b) is applicable to the situation being evaluated.

DEMAND: The time rate of energy flow. Demand usually refers to electric power and is measured in kW (equals kWh/h) but can also refer to natural gas, usually as Btu/hr, kBtu/hr, therms/day or ccf/day.

DEMAND (Utility): The rate or level at which electricity or natural gas is delivered to users at a given point in time. Electric demand is expressed in kilowatts (kW). Demand should not be confused with load, which is the amount of power delivered or required at any specified point or points on a system.

DEMAND BILLING: The electric capacity requirement for which a large user pays. It may be based on the customer's peak demand during the contract year, on a previous maximum or on an agreed minimum. Demand billing is measured in kilowatts.

DEMAND CHARGE: The sum to be paid by a large electricity consumer for its peak usage level.

DEMAND RESPONSIVENESS: Also sometimes referred to as load shifting. Activities or equipment that induce consumers to use energy at different (lower cost) times of day or to interrupt energy use for certain equipment temporarily, usually in direct response to a price signal. Examples include interruptible rates, doing laundry after 7 p.m., and air conditioner recycling programs.

DEMAND SAVINGS: The reduction in the demand from the pre-retrofit baseline to the post-retrofit demand, once independent variables (such as weather or occupancy) have been adjusted for. This term is usually applied to billing demand, to calculate cost savings or to peak demand, for equipment sizing purposes.

DEMAND SIDE MANAGEMENT (DSM): The methods used to manage energy demand including energy efficiency, load management, fuel substitution and load building. See LOAD MANAGEMENT.

DIRECT ENERGY SAVINGS (DIRECT PROGRAM ENERGY SAVINGS): The use of the words "direct savings" or "direct program savings" refers to the savings from programs that are responsible for the achievement of specific energy efficiency goals. Typically these are thought of as resource acquisition programs or programs that install or expedite the installation of energy-efficient equipment and which directly cause or help to cause energy efficiency to be achieved. Rebate, incentive or direct install programs provide direct energy savings.

DIRECT INSTALL or DIRECT INSTALLATION PROGRAMS: These types of programs provide free energy efficiency measures and their installation for qualified customers. Typical measures distributed by these programs include low flow showerheads and compact fluorescent bulbs.

DISTRIBUTED GENERATION: A distributed generation system involves small amounts of generation located on a utility's distribution system for the purpose of meeting local (substation level) peak loads and/or displacing the need to build additional (or upgrade) local distribution lines.

EFFICIENCY: The ratio of the useful energy delivered by a dynamic system (such as a machine, engine or motor) to the energy supplied to it over the same period or cycle of operation. The ratio is usually determined under specific test conditions.

EM&V: Evaluation, Measurement, Monitoring and Verification.

END-USE (MEASURES/GROUPS): Refers to a broad or sometimes narrower category that the program is concentrating efforts upon. Examples of end-uses include refrigeration, food service, HVAC, appliances, envelope and lighting.

ENERGY CONSUMPTION: The amount of energy consumed in the form in which it is acquired by the user. The term excludes electrical generation and distribution losses.

ENERGY COST: The total cost for energy, including such charges as base charges, demand charges, customer charges, power factor charges and miscellaneous charges.

ENERGY EFFICIENCY: Using less energy to perform the same function. Programs designed to use energy more efficiently - doing the same with less. For the purpose of this paper, energy efficiency programs are distinguished from DSM programs in that the latter are utility-sponsored and financed, while the former is a broader term not limited to any particular sponsor or funding source. "Energy conservation" is a term that has also been used but it has the connotation of doing without in order to save energy rather than using less energy to perform the same function and so is not used as much today. Many people use these terms interchangeably.

ENERGY EFFICIENCY IMPROVEMENT: Reduced energy use for a comparable level of service, resulting from the installation of an energy efficiency measure or the adoption of an energy efficiency practice. Level of service may be expressed in such ways as the volume of a refrigerator, temperature levels, and production output of a manufacturing facility or lighting level/square foot.

ENERGY EFFICIENCY MEASURE: Installation of equipment, subsystems or systems, or modification of equipment, subsystems, systems or operations on the customer side of the meter, for the purpose of reducing energy and/or demand (and, hence, energy and/or demand costs) at a comparable level of service.

ENERGY EFFICIENCY OF A MEASURE: A measure of the energy used to provide a specific service or to accomplish a specific amount of work (e.g., kWh/cubic foot of a refrigerator, therms/gallon of hot water).

ENERGY EFFICIENCY OF EQUIPMENT: The percentage of gross energy input that is realized as useful energy output of a piece of equipment.

ENERGY EFFICIENCY PRACTICE: The use of high-efficiency products, services and practices or an energy-using appliance or piece of equipment, to reduce energy usage while maintaining a comparable level of service when installed or applied on the customer side of the meter. Energy efficiency activities typically require permanent replacement of energy-using equipment with more efficient models. Examples: refrigerator replacement, light fixture replacement, cooling equipment upgrades.

ENERGY MANAGEMENT SYSTEM: A control system (often computerized) designed to regulate the energy consumption of a building by controlling the operation of energy consuming systems, such as the heating, ventilation and air conditioning (HVAC), lighting and water heating systems.

ENERGY SAVINGS: The reduction in use of energy from the pre-retrofit baseline to the post-retrofit energy use, once independent variables (such as weather or occupancy) have been adjusted for.

ENGINEERING APPROACHES: Methods using engineering algorithms or models to estimate energy and/or demand use.

ENGINEERING MODEL: Engineering equations used to calculate energy usage and savings. These models are usually based on a quantitative description of physical processes that transform delivered energy into useful work such as heat, lighting, or motor drive. In practice, these models may be reduced to simple equations in spreadsheets that calculate energy usage or savings as a function of measurable attributes of customers, facilities, or equipment (e.g., lighting use = watts × hours of use).

EVALUATION: The performance of studies and activities aimed at determining the effects of a program; any of a wide range of assessment activities associated with understanding or documenting program performance or potential performance, assessing program or program related markets and market operations; any of a wide range of evaluative efforts including assessing program-induced changes in energy efficiency markets, levels of demand or energy savings and program cost-effectiveness.

EX-ANTE SAVINGS ESTIMATE: Administrator-forecasted savings used for program and portfolio planning purposes as filed with the PA PUC, from the Latin for "beforehand."

EX-POST EVALUATION ESTIMATED SAVINGS: Savings estimates reported by the independent evaluator after the energy impact evaluation and the associated M&V efforts have been completed. If only the term "ex-post savings" is used, it will be assumed that it is referring to the ex-post evaluation estimate, the most common usage, from the Latin for "from something done afterward."

EX-POST (PROGRAM) ADMINISTRATOR-ESTIMATED SAVINGS: Savings estimates reported by the Administrator after program implementation has begun (Administrator-reported ex post), from the Latin for "from something done afterward."

EX-POST (PROGRAM) ADMINISTRATOR-FORECASTED SAVINGS: Savings estimates forecasted by the Administrator during the program and portfolio planning process, from the Latin for "from something done afterward."

EXTERNAL VALIDITY: The extent to which the association between an independent variable and a dependent variable that is demonstrated within a research setting also holds true in the general environment.

FREE-DRIVER: A non-participant who adopted a particular efficiency measure or practice as a result of a utility program. See SPILLOVER EFFECTS for aggregate impacts.

FREE-RIDER: A program participant who would have implemented the program measure or practice in the absence of the program.

GROSS SAVINGS: The change in energy consumption and/or demand that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated.

HETEROSCEDASTICITY: Unequal error variance. In statistics, a sequence or a vector of random variables is heteroscedastic if the random variables in the sequence or vector may have different variances. This violates the regression assumption of constant variance (the variance of the errors is constant across observations or homoscedastic). Typically, residuals are plotted to assess this assumption. Standard estimation methods are inefficient when the errors are heteroscedastic. A common example is when variance is expected to be greater on a variable measurement for larger firms than for smaller firms.

HOMOSCEDASTIC (HOMOSCEDASTICITY): Constant error variance, an assumption of classical regression analysis. See also HETEROSCEDASTICITY.

IMPACT EVALUATION: Used to measure the program-specific induced changes in energy and/or demand usage (such kWh, kW and therms) and/or behavior attributed to energy efficiency and demand response programs.

IMPACT YEAR: Depending on the context, impact year means either (a) the twelve months subsequent to program participation used to represent program costs or load impacts occurring in that year, or (b) any calendar year after the program year in which impacts may occur.

INCENTIVES: Financial support (e.g., rebates, low-interest loans) to install energy efficiency measures. The incentives are solicited by the customer and based on the customer's billing history and/or customer-specific information.

INDEPENDENT VARIABLES: The factors that affect the energy and demand used in a building but cannot be controlled (e.g., weather or occupancy).

INDIRECT ENERGY SAVINGS (INDIRECT PROGRAM ENERGY SAVINGS): The use of the words "indirect savings" or "indirect program savings" refers to programs that are typically information, education, marketing or outreach programs in which the program's actions are expected to result in energy savings achieved through the actions of the customers exposed to the program's efforts, without direct enrollment in a program that has energy savings goals.

LOAD SHAPES: Representations such as graphs, tables, and databases that describe energy consumption rates as a function of another variable such as time or outdoor air temperature.

INTERNAL VALIDITY: The validity of (causal) inferences in scientific studies, usually based on experiments as experimental validity. Inferences are said to possess internal validity if a causal relation between two variables is properly demonstrated.

MARKET EFFECT EVALUATION: The evaluation of the change in the structure/functioning of a market or the behavior of participants in a market that results from one or more program efforts. Typically the resultant market or behavior change leads to an increase in the adoption of energy-efficient products, services, or practices.

MARKET TRANSFORMATION: A reduction in market barriers resulting from a market intervention, as evidenced by a set of market effects, that lasts after the intervention has been withdrawn, reduced, or changed.

MEASUREMENT: A procedure for assigning a number to an observed object or event.

MEASUREMENT AND VERIFICATION (M&V): Data collection, monitoring, and analysis associated with the calculation of gross energy and demand savings from individual sites or projects. M&V can be a subset of program impact evaluation.

MEASUREMENT BOUNDARY: The boundary of the analysis for determining direct energy and/or demand savings.

METERING: Meeting is the collection of energy consumption data, over time, through the use of meters. These meters may collect information with respect to an end-use, a circuit, a piece of equipment, or a whole building (or facility). Short-term metering generally refers to data collection for no more than a few weeks. End-use metering refers specifically to separate data collection for one or more end-uses in a facility, such as lighting, air conditioning or refrigeration. Spot metering is an instantaneous measurement (rather than over time) to determine an energy consumption rate.

MONITORING: Gathering of relevant measurement data, including but not limited to, energy consumption data over time to evaluate equipment or system performance, e.g., chiller electric demand, inlet evaporator temperature and flow, outlet evaporator temperature, condenser inlet temperature, and ambient dry-bulb temperature and relative humidity or wet-bulb temperature, for use in developing a chiller performance map (e.g., kW/ton vs. cooling load and vs. condenser inlet temperature).

MULTI-COLINEARITY: A statistical phenomenon in which two or more predictor variables in a multiple regression model are highly correlated. In this situation the coefficient estimates may change erratically in response to small changes in the model or the data. Multi-Colinearity does not reduce the predictive power or reliability of the model as a whole, at least within the sample data themselves; it only affects calculations regarding individual predictors.

NET SAVINGS: The total change in load that is attributable to an energy efficiency program. This change in load may include, implicitly or explicitly, the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, participant and non-participant spillover and other causes of changes in energy consumption or demand.

NET-TO-GROSS RATIO(NTGR): A factor representing net program savings divided by gross program savings that is applied to gross program impacts to convert them into net program load impacts.

NON-PARTICIPANT: Any consumer who was eligible, but did not participate in the subject efficiency program in a given program year. Each evaluation plan should provide a definition of a non-participant as it applies to a specific evaluation.

NON-RESPONSE BIAS: The effect of a set of respondents refusing or choosing not to participate in research; typically larger for self-administered or mail-out surveys.

NORMALIZED ANNUAL CONSUMPTION (NAC) ANALYSIS: A regression-based method that analyzes monthly energy consumption data.

PARTIAL FREE-RIDER: A program participant who would have implemented, to some degree, the program measure or practice in the absence of the program (i.e., a participant may have purchased an ENERGY STAR appliance in the absence of the program, but because of the program the participant purchases an appliance that is higher in efficiency).

PARTICIPANT: A consumer that received a service offered through the subject efficiency program, in a given program year. The term "service" is used in this definition to suggest that the service can be a wide variety of services, including financial rebates, technical assistance, product installations, training, energy efficiency information or other services, items, or conditions. Each evaluation plan should define "participant" as it applies to the specific evaluation.

PEAK DEMAND: The maximum level of metered demand during a specified period, such as a billing month or a peak demand period.

PERSISTENCE STUDY: A study to assess changes in program impacts over time (including retention and degradation).

PORTFOLIO: Either (a) a collection of similar programs addressing the same market (e.g., a portfolio of residential programs), technology (e.g., motor efficiency programs), or mechanisms (e.g., loan programs) or (b) the set of all programs conducted by one organization, such as a utility (and which could include programs that cover multiple markets, technologies, etc.).

PRECISION: The indication of the closeness of agreement among repeated measurements of the same physical quantity.

PROCESS EVALUATION: A systematic assessment of an energy efficiency program for the purposes of documenting program operations at the time of the examination, and identifying and recommending improvements to increase the program's efficiency or effectiveness for acquiring energy resources while maintaining high levels of participant satisfaction.

PROGRAM: A group of projects, with similar characteristics and installed in similar applications. Examples could include a utility program to install energy-efficient lighting in commercial buildings, a developer's program to build a subdivision of homes that have photovoltaic systems, or a state residential energy efficiency code program.

PROJECT: An activity or course of action involving one or multiple energy efficiency measures, at a single facility or site.

REGRESSION ANALYSIS: Analysis of the relationship between a dependent variable (response variable) to specified independent variables (explanatory variables). The mathematical model of their relationship is the regression equation.

RELIABILITY: Refers to the likelihood that the observations can be replicated.

REPORTING PERIOD: The time following implementation of an energy efficiency activity during which savings are to be determined.

RETROFIT ISOLATION: The savings measurement approach defined in IPMVP Options A and B, and ASHRAE Guideline 14, that determines energy or demand savings through the use of meters to isolate the energy flows for the system(s) under consideration.

RIGOR: The level of expected confidence and precision. The higher the level of rigor, the more confident one is that the results of the evaluation are both accurate and precise.

SPILLOVER: Reductions in energy consumption and/or demand caused by the presence of the energy efficiency program, beyond the program-related gross savings of the participants. There can be participant and/or nonparticipant spillover.

STATISTICALLY ADJUSTED ENGINEERING (SAE) MODELS: A category of statistical analysis models that incorporate the engineering estimate of savings as a dependent variable.

STIPULATED VALUES: See "deemed savings."

UNCERTAINTY: The range or interval of doubt surrounding a measured or calculated value within which the true value is expected to fall within some degree of confidence.

VALUE OF INFORMATION: A balance between the level of detail (rigor) and the level of effort required (cost) in an impact evaluation.

Appendix B. Example of Standardized Method for Determining Free-Ridership Rates

PA Consulting prepared a report in June 2003 for National Grid, NSTAR Electric, Northeast Utilities, Unitil, and Cape Light Compact that provides standardized methods for the measurement of free-ridership within commercial and industrial energy efficiency programs in Massachusetts. The research approach recommended by PA Consulting in their June 2003 report can be used for determining free-ridership in residential, commercial or industrial programs. *Note: Each reference to "company" within the following questionnaire has been re-worded to reflect a survey directed towards an individual for use within any energy efficiency program.*

Table 1: Free-Rider Questionnaire – Initial Measure

Initial Measure-Specific Free-Rider Questions	
According to our records, the total cost for all [measure # description] installed was about [measure # total project cost]. [Sponsor] paid about [measure # sponsor contribution] of the total cost of this equipment. [NON-TECHNICAL ASSESSMENT: You might also have received some technical assistance from a [sponsor] rep, engineer, or equipment vendor.] [TECHNICAL ASSESSMENT: The program also contributed toward the cost of a Technical Assessment Study.]	
F1	If [sponsor] had not paid a portion of the equipment cost or provided any technical assistance or education through the [program], would you have purchased any [measure # description] within one year of when it was installed? **(TIMING)** (IF NO, SKIP TO CONSISTENCY QUESTIONS)
F2	Without the program [contribution/incentive/rebate], technical assistance, or education, would you have purchased the exact same quantity of [measure # description] within one year? **(QUANTITY)**
F2b	(IF NO) What percent of this [measure # description] do you think you would have purchased on its own within one year?
F3	You said you would have installed at least some of [measure # description] on your own if the program had not been available. What percent of this equipment would have been of the same efficiency or higher efficiency as what was installed through the program? **(PROGRAM EFFICIENCY – NOT RELEVANT TO CFLs)**

Initial Measure-Specific Free-Rider Questions	
F4	Now I want to focus on what it would have cost you to install this equipment on its own without the program. Do you think you would have paid the additional [measure # sponsor contribution], on top of the amount you already paid, to install the same quantity and efficiency of [measure # description] within one year? **(COST)**
F5	(IF NO) How would you have adjusted your purchase to accommodate the fact that you wouldn't have paid all of the costs? Would you have purchased less equipment, lower efficiency equipment, or done something else? **(ELASTICITY)**

Table 2: Free-Rider Questionnaire – Consistency Check Questions

Consistency Checks		Free-ridership Rate		
		0%	100%	
P3	Did you have specific plans set aside to install any of this equipment before you talked with anyone about the program? **(INTENT)** (If YES) What plans Existed? (Probe for timing, quantity, and efficiency)	Yes	No	Yes
P4	(If P3=YES) Was it necessary to change the type or efficiency level of equipment in your plans in order to qualify for the [program]? (IF YES) What changes were made? (Probe for timing, quantity, and efficiency)	No	Not Asked	Yes
F8	Was the information or advice you received from a contractor, design team, utility rep, or an engineer a crucial factor in your decision to install this high efficiency equipment through the program at the time you did?	No	Yes	
F9	(OPEN ENDED) I'd like to better understand your purchase decision. Maybe you could just describe in your own words what impact, if any, the program had on your decision to install the energy efficient [measure # description] at the time you did?	ASKED	ASKED	

The above survey approach has been used by PA Consulting for the state of Massachusetts to determine free-ridership within commercial and industrial energy efficiency programs.

Appendix C. Sampling Roadmaps

C.1. Stratified Sampling Roadmap

Step 1: Create a database listing each project in the population and the corresponding tracked savings estimates.

- The tracked values are designated by variable x_i. *i* is the project, *i*=1, …, N, and N is the number of projects in the population.

Step 2: Sort the list of projects by increasing standard deviation.

a) Estimate the expected error ratio, *er*. It the tracking system is expected to provide accurate estimates of the actual savings of most sample projects, then the *er* is likely to be relatively small, e.g., near 0.4. On the other hand, if the tracking system is expected to provide poor estimates of the actual savings, then the *er* should be closer to 1.0.

b) Calculate the standard deviation.

c) Calculate the assumed secondary equation of the ratio model for each project according to the following equation: $\sigma_i = \sigma_0 x_i^{\gamma}$. Assume γ=0.8.

d) Sort the list by increasing σ_i.

e) Calculate the cumulative sum of the σ_i, $c_i = \sum_{j=1}^{i} \sigma_j$.

Step 3: Group projects into strata.

a) Choose a desired number of strata *L* (usually three to five).

b) Divide the projects in the sorted list into *L* strata, designated by the variable *h*, so that the cumulative sum of each strata is approximately equal. In Excel, this can be done by the following formula: $h_i = INT\left(L \frac{c_i}{c_N + 0.999} \right)$.

Step 4: Select sample projects to survey.

a) Calculate the overall sample size according to the random sample protocol according to the specified level of confidence and precision and assuming the same error ratio used in Step 2 as the coefficient of variance.

b) Allocate the ample equally to each stratum. If the sample size in a particular stratum exceeds the population size of that stratum, all projects should be sampled in that stratum. And, if desired, the sample may be increased in the remaining strata so that the overall population sample size is closer to the planned value.

Assuming the population was appropriately stratified, sampled and surveyed, the following characteristics can be extrapolated from the data collected to represent the entire population. The stratified ratio estimate method is used to define the relationship between a measurable variable and a second variable; for the impact evaluation the measurable variable, *y*, will be the measured savings while the second variable, *x*, will be the tracked or expected savings. As noted earlier, one of the key population parameters in this method is the ratio between actual measured savings and tracked savings. Using the data collected from a sample of the projects, the ratio estimator, *b*, can be estimated as follows:

$$b = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i x_i}$$ **Equation C-1**

In the above equation, *n* is the sample size and *w* is the case weight. The case weight is defined according to the following equation where N_h is the total number of population projects in stratum *h* and n_h is the number of projects randomly selected from stratum *h*:

$$w_i = \frac{N_h}{n_h}$$ **Equation C-2**

The statistical precision of *b* can be assessed by calculating the standard error using the following equation:

$$se(b) = \frac{1}{\hat{X}} \sqrt{\sum_{i=1}^{n} w_i (w_i - 1) e_i^2}$$ **Equation C-3**

Here $\hat{X} = \sum_{i=1}^{n} w_i x_i$ and $e = y_i - bx_i$. The error bound and relative precision can then be calculated according to the following equations, using the z-value of the prescribed confidence level of the study.

$$eb(b) = z \cdot se(b)$$ **Equation C-4**

The data collected from the survey can also be used to estimate the total actual savings of the population of projects. The equations used to estimate the total savings and the corresponding standard error are found below.

This section provides a roadmap for determining sample size based on the following expectations, assumptions, and program characteristics:

- Expected confidence and precision level
- Assumed error ratio or coefficient of variance
- Total projects within a given program
- Desired sample design

The following figure outlines the general steps used to design and select and sample for surveying.

$$rp = \frac{eb(b)}{b}$$ **Equation C-5**

$$\hat{Y} = b\sum_{i=1}^{N} x_i$$ **Equation C-6**

$$se(\hat{Y}) = se(b)\sum_{i=1}^{N} x_i$$ **Equation C-7**

Appendix D. Audit Activities Checklist

Audit Activities Checklist: Checklist of activities to be completed by the Statewide Evaluation Team for program audits and field inspections.

EDC: [EDC Name] **Program:** [Program Title]

Item 1: Define measure type and program characteristics

☐ Deemed ☐ Partially Deemed ☐ Custom

Item 2: Determine VOI

Anticipated Energy savings: [Energy Savings]
Anticipated Demand savings: [Demand Savings]
Anticipated projects: [Anticipated No. Projects]
Priority: ☐ High ☐ Medium ☐ Low

Item 3: EDC EM&V Plan Review

☐ Review measure categorization
☐ Review M&V protocols
☐ Review implementation plan
☐ Review sample design
☐ Review data for impact evaluation
☐ Review data for TRC calculation
☐ Check for compliance against
☐ Miscellaneous

Item 4: EDC EM&V Plan Implementation

☐ Schedule
☐ Data collection
☐ Sample design
☐ Site-inspections
☐ Measure verification
☐ Uncertainty documentation
☐ Pre-project applications
☐ Data quality
☐ Miscellaneous Category

Item 5: SWE Audit Activities

☐ Independent Site-Visits [Spot-Checks]
☐ Joint Impact Evaluations
☐ QC data input
☐ Review calculations
☐ Review Pre-applications
☐ Verify data collected
☐ Inspect instrumentation
☐ Review models and simulations
☐ Review uncertainty documentation
☐ Review statistical summaries

Item 6: SWE Evaluation of Impact Reports

☐ By project
☐ By sample
☐ By program
☐ By portfolio

Item 7: Audit Report

☐ Quarterly
☐ Annual

Item 1: Define measure type and program characteristics

☐ Deemed ☐ Partially Deemed ☐ Custom

[Insert program description, unique characteristics, and key factors here.]

Item 2: Determine VOI

Priority: ☐High ☐Medium ☐Low

- Anticipated energy savings: [Energy Savings]
- Anticipated demand savings: [Demand Savings]
- Anticipated projects: [No. Projects]
- Anticipated participants: [No. Participants]
- Anticipated savings per project: [Savings per Project]
- Anticipated savings per participant: [Savings per Participant]

[Insert information regarding VOI judgment and justification for evaluation and audit activities.]

Item 3: EDC EM&V Plan Review

The EM&V Plan review will verify that the plans incorporate the appropriate protocols as they are described in (a) the Audit Plan, (b) the TRM Order, (c) the TRC Order, and (d) any supplemental memos or protocols accepted by all parties. The SWE will specifically check for compliance in the following categories[71] (check all that apply)*:*

- ☐ Measure categorization:
- ☐ M&V protocols
  - o IPMVP Option
  - o Custom Protocols
  - o Supplemental Protocols
  - o Combination of Protocols
- ☐ Sample design
  - o Confidence Level
  - o Precision Range
- ☐ Implementation plan
  - o Schedule
  - o Training exercises
- ☐ Survey or other instruments
  - o Participant/Non-participant Surveys
  - o Engineering instruments
- ☐ Data collected for impact evaluation
- ☐ Data collected for TRC calculation
- ☐ Additional categories: [Insert additional categories here.]

M&V Protocol Option:

[Description of approved protocol option.]

Sample Design:

[Description of approved sample design.]

[71] Note: this list is not exclusive and other categories may be added as appropriate for the particular program.

Data to be collected:

SITE INFORMATION	REQUIRED DOCUMENTATION
[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]
EXISTING EQUIPMENT INFORMATION	
[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]
NEW EQUIPMENT INFORMATION	
[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]

Notes: [Insert any additional comments that might pertain to the review of the EM&V Plan for this particular program.]

Item 4: EDC EM&V Plan Implementation

The audit of the EM&V Plan implementation is the next stage of the audit process. Once the EM&V plans are approved, it will be paramount that SWE Team verify that the actual field activities and engineering calculations laid out in the plan be implemented accordingly. The activities undertaken by the SWE in this portion of the audit serve to verify implementation of the EM&V. These activities will not be used to validate any assumptions, calculations, measurements, etc. Thus, all activities are either a review or observation of an event carried out by the EDC or EDC evaluator. The SWE will specifically check for compliance in the following categories[72] (check all that apply)*:*

	CATEGORY:	METHOD*:	SAMPLE**:
☐	SCHEDULE		
☐	DATA COLLECTION[73]		
☐	SAMPLE DESIGN[74]		
☐	SITE-INSPECTIONS		
☐	MEASURE VERIFICATION		
☐	MEASURE INSTALLATION QUALITY		
☐	UNCERTAINTY DOCUMENTATION		
☐	PRE-PROJECT APPLICATIONS		
☐	DATA QUALITY[75]		
☐	ADDITIONAL CATEGORIES: [LIST HERE.]		

** Examples: Completed field surveys, database entries, Join Impact Evaluation observations, project applications/invoices, etc.*

***Examples: not applicable, all files, all fields, statistical sample of EDC sample, X amount, X%, etc.*

[72] Note: This list is not exclusive and other categories may be added as appropriate for the particular program.
[73] Note: Checking for completeness in data collection.
[74] Note: Verification of sample size, project selection, and issues pertaining to non-response.
[75] Note: Checking for data entry accuracy.

Data Collection:

☐	**SITE INFORMATION**	**REQUIRED DOCUMENTATION**
	[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]
☐	**EXISTING EQUIPMENT INFORMATION**	
	[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]
☐	**NEW EQUIPMENT INFORMATION**	
	[INSERT REQUIREMENTS.]	[INSERT REQUIREMENTS.]

Notes: [Insert any additional comments that might pertain to the implementation audit of the EM&V Plan for this particular program.]

Item 5: Independent SWE Audit Activities

The audit of the EM&V Plan implementation is the next stage of the audit process. Once the EM&V plans are approved, it will be paramount that SWE Team validate the actual field measurements and engineering calculations. The activities undertaken by the SWE in this portion of the audit serve to validate the numerical findings of the EM&V activities and impact evaluations. The findings from this audit will be used to assess the validity of assumptions, calculations, measurements, etc. Thus, all activities require an actionable step by the SWE in order to assess the corresponding findings of the EDC or EDC evaluator. The SWE will specifically check for validity in the following categories[76] (check all that apply)*:*

- ☐ Measure energy consumption
- ☐ Measure operating hours
- ☐ Measure performance parameters conditions
- ☐ Measure mechanical specifications
- ☐ Measure installation quality
- ☐ Site-specific measurements
- ☐ External factors affecting savings (i.e., weather, occupancy, etc.)
- ☐ No. measure installed on-site
- ☐ Calculation of savings
  - ☐ Engineering calculations
  - ☐ Simulations
  - ☐ Regression analysis
  - ☐ Additional Methods: [List here.]
- ☐ Other: [List here.]

The SWE will check for validity in EDC evaluation findings by conducting the following activities listed on the next page (check all that apply):

[76] Note: This list is not exclusive and other categories may be added as appropriate for the particular program.

Potential Audit Activities based of M&V Protocols Specified for Program.

Audit Category		Audit Activities	Method*	Sample**
Pre-retrofit Energy Measure	☐	• Verify that the isolation metering reflects the boundary between equipment which the ECM affects and that which it does not affect.		
	☐	• Verify measurement of all open variables as defined in M&V Plan.		
	☐	• Measure all open variables as defined in M&V Plan.		
	☐	• Verify measurement of all variables as defined by the TRM		
	☐	• Measure all variables as defined by the TRM		
	☐	• Verify that meter or billing data reflects the energy use associated with the building in which the ECM was installed.		
	☐	• Verify data collection of all applicable meters and/or sub-meters.		
	☐	• Collect data from all applicable meters and/or sub-meters		
Verification of Installation	☐	• Verify that conditions of project conform to guidelines of program (i.e., verify eligibility).		
	☐	• Verify measure existence.		
	☐	• Count installations.		
	☐	• Verify quality of installation.		
	☐	• Verify the occurrence of re-inspections post-retrofit regarding ECM performance and operating conditions.		
Post-retrofit Energy Measure	☐	• Verify that the isolation metering reflects the boundary between equipment which the ECM affects and that which it does not affect.		
	☐	• Verify measurement of open variables as defined in M&V Plan.		
	☐	• Measure open variables as defined in M&V Plan.		
	☐	• Verify frequency of measurement adheres to M&V protocols (e.g., intermittent, short-term, or continuous).		
	☐	• Verify measurement of all variables as defined by the TRM		
	☐	• Measure all variables as defined by the TRM		
	☐	• Verify that meter or billing data reflects the energy use associated with the building in which the ECM was installed.		
	☐	• Verify data collection of all applicable meters and/or sub-meters.		
	☐	• Collect data at all applicable meters and/or sub-meters.		
	☐	• Verify frequency of measurement adheres to M&V protocols.		
	☐	• Collected data with frequency outlined in M&V protocols.		
On-Site ECM Sampling	☐	• Verify on-site ECM sampling design adheres to protocols.		
	☐	• Review on-site ECM sampling calculations.		
	☐	• Re-calculate ECM sample.		
	☐	• Verify M&V of ECM sample.		
Uncertainties/Independent Variables	☐	• Verify documentation of uncertainties.		
	☐	• Document uncertainties.		
	☐	• Review documented uncertainties for accuracy or plausibility.		
	☐	• Verify documentation of independent variable conditions.		
	☐	• Document independent variable conditions.		

	☐	• Review documented independent variable conditions for accuracy or plausibility.		
Instrumentation	☐	• Verify equipment calibration according to NIST procedures.		
	☐	• Calibrate equipment according to NIST procedures.		
	☐	• Verify latest measurement techniques according to IPMVP are implemented.		
	☐	• Verify that protocols for data collection errors and lost data are followed.		
Implementation	☐	• Verify the timeliness of participant notification.		
	☐	• Verify competency of field-staff.		
	☐	• Verify all protocols identified in M&V Plan are followed.		
	☐	• Follow all M&V protocols specified in EM&V Plan.		
Data Tracking and Reporting	☐	• Verify all variables accounted for in tracking system.		
	☐	• QC data tracking and reporting system against EM&V Documents		
Perform Engineering Impact Calculations	☐	• Review project impact calculations.		
	☐	• Conduct independent project impact calculations.		
	☐	• Review base year model.		
	☐	• Create base year model.		
	☐	• Assess statistical validity of selected model.		
Extrapolation of Sample Findings to Impact	☐	• Review sample statistics calculations.		
	☐	• Conduct independent sample statistics calculations.		

** Examples: Completed field surveys, database entries, Join Impact Evaluation observations, project applications/invoices, etc.*

***Examples: not applicable, all files, all fields, statistical sample of EDC sample, X amount, X%, etc.*

Notes: [Insert any additional comments that might pertain to the validation of M&V and evaluation findings for this particular program or the corresponding Audit Activities to be performed by the SWE Team.]

Item 6: SWE Evaluation of Impact Reports

The SWE will assess the impact reports at various levels. The assessment will be based on the findings of the previous two audit steps (EDC EM&V Plan Implementation and SWE Audit Activities). The potential for SWE assessment include (check all that apply):

- ☐ Project savings
- ☐ Participant savings
- ☐ EDC Sample savings
- ☐ Program realization rates
- ☐ Program impacts
- ☐ Portfolio impacts
- ☐ Other: [List here.]

Notes: [Insert any additional comments that might pertain to the SWE evaluation if impact reports.]

Item 7: Audit Report

☐ Quarterly
☐ Annual
Final Approval Status:
☐ Rejected ☐ Partially Approved ☐ Approved

The SWE Team will submit **quarterly reports** to the Commission with updates on:

☐ Impact evaluations
☐ Cost-effectiveness
☐ Process evaluations

These reports will identify:

☐ Best practices exhibited to date
☐ Areas for improvements
☐ Recommendation for updating targets based on the current findings

The SWE Team will verify and support with confidence the gross savings for program year one and the gross and net savings for subsequent program years in annual reports. The annual reports will include the following data for each program and combinations of programs so that information is reported by end-user market sector (residential, low-income residential, commercial/industrial, and governmental, education, and non-profit entities):

☐ Annual kWh savings,
☐ On peak demand kW reduction coincident with the summer peak
☐ Cost-effectiveness according to the TRC test
☐ Levelized cost per lifetime kWh saved

The reports will include savings estimates:

☐ For each program year and
☐ Cumulatively, for all program years.

Savings estimates will be reported as a point estimate with upper and lower bounds to designate the range according to EM&V confidence levels as addressed in the Audit Plan.

The annual reports will also contain a complete description of the audit plan including the audit activities and findings to date. The reports will contain all documentation reasonably needed to follow the analysis starting with each measure and project in the sample up to the final realization rates and their application to each EDC's reported savings.

The following is a preliminary outline for the annual evaluation report:

- An executive summary.
- A discussion of the methodology used in the evaluation and audit, highlighting any changes to the process program year-to-year.
- A description of the data-collection, matching, and cleaning process.
- A detailed discussion of all the findings (verified on the audits), including the total energy and demand impacts (realization rate) by EDC, program, and sector.
- An assessment of tracking and calculation of energy savings, customer satisfaction, trade-ally satisfaction, educational components, and missed opportunities analysis.
- An assessment of each EDC's plan to determine whether EDCs are meeting energy savings and load reduction targets.
- A review of each EDC's M&V plans and execution of said plans, with a list of recommendations and updates as necessary.
- A conclusion addressing the status of the overall portfolio of programs in terms of meeting the energy and demand savings targets and the cost-effectiveness of the programs offered.

Note: All project and EDC specific information will be kept confidential where necessary.

Notes: [Insert any additional comments that might pertain to the SWE Audit Reports.]

Appendix E. Impact Evaluation M&V Data Requirements

E.1. Commercial and Industrial Measures M&V Data Requirements

The following sections outline the M&V data requirements by C&I program type.

E.1.1. Existing Facility Lighting Efficiency Programs

Example Measures: Prescriptive Lighting, Fluorescent Lighting (Super T-8) and Custom Lighting

The recommendations for the M&V plan described here have two purposes:

- To create an audit trail for the SWE, thereby enabling inspection of facilities to accurately account for hours of use of the installed fixtures
- To determine and have record of the preexisting condition as may be required.
- To enable the EDCs and the SWE to determine the actual savings.

Lighting contractors and Conservation Service Providers (CSPs) who are responsible for developing and implementing lighting projects are responsible for providing lighting audit information including pre- and post-fixture types and quantities by area and an estimation of hours of use for each fixture or fixture group. This information should be part of the application process and confirmed by the CSP. The EDC's evaluator, with oversight by the SWE, will subsequently sample these project files and conduct post installation site inspections. Logging of lighting hours of use may be required at the discretion of the SWE for individual projects sampled by the EDC evaluator. Collaboration in sampling protocol, site visits and analysis will be required between the EDC evaluator and the SWE to optimize evaluation costs and provide independent confirmation of realization rates. The CSP should be required by the utilities to maintain complete project files for all participants. The information is necessary to provide accountability and more accurately assess the gross energy and demand savings of the programs.

Program managers should create a standardized lighting audit and application template to document the required information. This will help maintain the uniformity and organization of the documentation throughout each project. It should be the responsibility of the program managers to ensure that contractors and CSPs record the necessary information and that it is complete.

Table E-1: C&I Lighting Program Information to Be Collected

Site Information	
Name of customer Address Phone number Standard building type Utility and account number (electric)	Statement on Application Form
Existing Equipment Information	**Required Documentation**
Lighting Audit Pre Condition Lamp Ballast Wattage from Approved Table When existing equipment was installed (year) or if this is a new installation.	Lighting Audit Form, Application Stage
Line Item Quantity	
Line item Hours of Use	
New Equipment Information	**Required Documentation**
Lighting Audit Post Condition Lamp Ballast Wattage from Approved Table	Lighting Audit, As Built
Line Item Quantity	CSP Inspection, Acceptance of As Built
Line item Hours of Use	CSP Inspection, Acceptance of As Built
Equipment & Installation Costs	

E.1.2. High Efficiency Electric Motors

Example Measures: High efficiency electric motors purchased through participating suppliers that replace standard efficiency electric motors.

This measure covers new, general-purpose, three-phase induction motors that meet the following criteria:

- NEMA Design A or B
- From 1 to 200 rated horsepower
- Totally enclosed fan cooled (TEFC) or open drip proof (ODP)
- Perform equal to or better than the NEMA Premium efficiency

Program managers should ensure that the motor end user or motor installation contractor record all information on the new motor nameplate and the motor being replaced and collect a copy of the purchase order or invoice for the new motor. This nameplate information will be required to accurately assess the gross energy and demand savings. It may be necessary to measure load versus name plate to determine load factor.

Program managers should create a standard format to document the required information. This will help maintain the uniformity and organization of the documentation. It should be the responsibility of the program managers to ensure contractors record the necessary information and that it is complete.

Table E-2: C&I Motor Program Information to Be Collected

Site Information	Required Documentation
Name of customer Address Phone number Standard building type Electric Utility and Account Number	Statements on application form
Existing Motor Information	
Motor nameplate information, including manufacturer, model number, rated horsepower, enclosure type, speed, NEMA nominal efficiency Description of Application Estimated annual operating hours Constant or variable speed When existing equipment was installed (year) of if this is a new installation	Statements on application form Photograph Obtain "Run Load Factor"
New Motor Information	
Motor nameplate information, including manufacturer, model number, rated horsepower, enclosure type, speed, and NEMA nominal efficiency Constant or variable speed	Statements on application form Vendor Specification Sheet Obtain "Run Load Factor"
Equipment & Installation Costs	

Inspections should be performed on a sample of installations.

E.1.3. HVAC Systems

Example Measures: Electric Chillers, Unitary HVAC/Split Systems, Air-to-Air Heat Pump Systems, Water Source Heat Pumps, Ground Source Heat Pumps and Packaged Terminal Systems

HVAC equipment will be provided through participating installation contractors and dealers who will also install the HVAC equipment. The installation contractors should be required by the utilities to collect and record information about each project that receives an incentive through the program. This information is necessary to provide accountability and more accurately assess the gross energy and demand savings of the programs.

In certain large commercial HVAC projects with multiple redundant systems, the verification of savings using Equivalent Full Load Hours (EFLH) may not appropriately characterize the operation and savings in practice. In these cases, based on a size threshold, an enhanced protocol may be recommended by the SWE. This issue will be considered in the TWG relative to the appropriate program level M&V protocol.

Program managers should create a standard format to document the required information. This will help maintain the uniformity and organization of the documentation throughout each project. It should be the responsibility of the program managers to ensure contractors record the necessary information and that it is complete.

Table E-3: C&I HVAC System Program Information to Be Collected

Site Information	Required Documentation
Name of customer Address Phone number Standard building type Electric Utility and Account Number	Statements on application form
Existing Equipment Information	
Equipment Number Type Manufacturer and model number Efficiency: EER, SEER, COP, kW/ton Cooling Capacity Heating Capacity (Heat Pumps) Quantity When existing equipment was installed (year) or if this is a new installation Estimated annual operating hours Year of manufacture Fuel switching?	Statements on application form Photograph
New Equipment Information	
Type Manufacturer and model number Efficiency: EER, SEER, COP, kW/ton Cooling Capacity Heating Capacity (Heat Pumps) Quantity Estimated annual operating hours Year of manufacture	Statements on application form Vendor Specification Sheet
Equipment & Installation Costs	

Inspections should be performed on a random sample of installations.

E.1.4. Variable Frequency Drive Programs

Measures: Variable Frequency Drive Programs (VFD)

The current TRM provides a measurement protocol for Variable Frequency Drives (VFDs). Central to calculating savings using the fan laws is the load shape of the flow rate of the fluid before the installation of the VFD and the load shape of the flow rate of the fluid after installation of the VFD. This defines the difference in work required and therefore is the basis of the energy savings.

The data required to verify savings and the steps required to implement the M&V plan will, to some degree, depend on the work of the TWG and changes to the TRM approved by the PA PUC. However the recommendations for the M&V plan described here has two additional purposes:

- To create an audit trail for the SWE enabling inspection of facilities to accurately account for load on the VFD as installed and to determine and have record of the pre-existing condition as may be required; and
- To enable the EDCs and the SWE to determine the actual savings versus the savings deemed, based on factors of savings per horsepower for air handling systems and chilled water systems, to inform discussions of proscriptive changes to the TRM.

Mechanical contractors and CSPs who are responsible for developing and implementing VFD projects are responsible for providing information sufficient to estimate the pre- and post-condition energy use and fluid flow rates over time. For example, this is usually done by either the CSP or EDC, depending on the program logic, before installation by spot measurement of kW on a constant speed motor to be replaced by the VFD and modeling of the expected flow rates taking into account the proposed control configuration and weather bin data. This information should be part of the application process and be conducted and/or confirmed by the CSP as appropriate.

In certain Custom Measure applications, pre- or post-installation metering of the load over a representative time period may be needed to verify savings when loading is variable. The EDC evaluator, with oversight by the SWE, will subsequently sample these project files and conduct post installation site inspections. Short term metering of VFD kW may be required at the discretion of the SWE for individual projects sampled by the EDC evaluator to confirm modeled data or metered data. Collaboration in sampling protocol, site visits and analysis will be required between the EDC evaluator and the SWE to optimize evaluation costs and provide independent confirmation of realization rates. The CSP should be required by the utilities to maintain complete project files for all participants. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Program managers should create a standardized file and application template to document the required information. This will help maintain the uniformity and organization of the documentation throughout each project. It should be the responsibility of the program managers to ensure that contractors and CSPs record the necessary information and that it is complete.

Table E-4: C&I Variable Frequency Drive Program Information to Be Collected

Site Information	
Name of customer Address Phone number Standard building type Utility and account number (electric)	Statement on Application Form
Existing Equipment Information	**Required Documentation**
Existing Motor Name Plate Data Pre kW , Instantaneous or Pre kW over Representative Loading Period When existing equipment was installed (year) or if this is a new installation	Motor Inventory Form, Application Stage Logged kW Data Modeled Savings
Line Item Quantity	Motor Inventory Form, Application Stage
Line item Hours of Use (Not FLH or EFLH)	Motor Inventory Form, Application Stage
New Equipment Information	**Required Documentation**
Proposed Motor Name Plate Data Drive Nameplate data and Efficiency, Post kW over Representative Loading Period	Motor Inventory Form, Application Stage Logged kW Data, post
Line Item Quantity	Motor Inventory Form, Application Stage
Line item Hours of Use (Not FLH or EFLH)	Motor Inventory Form, Application Stage
Equipment & Installation Costs	

E.2. Residential Measure M&V Data Requirements

The following sections outline the M&V data requirements by residential program type.

E.2.1. HVAC Efficiency Programs

Measures: Central Air Conditioner (A/C) and Air Source Heat Pump (ASHP), Ground Source Heat Pump (GSHP), GSHP De-superheater, and Furnace High Efficiency Fan

HVAC equipment will be provided through participating installation contractors and dealers who will also install the HVAC equipment. The installation contractors should be required by the utilities to collect and record information about each project that receives an incentive through the program. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Program managers should create a standardized format to document the required information. This will help maintain the uniformity and organization of the documentation throughout each project. It should be the responsibility of the program managers to ensure contractors record the necessary information and that it is complete.

Table E-5: Residential HVAC Efficiency Program Information to Be Collected

Site Information	
Name of customer Address Phone number Utility and account number (electric)	Statement on Application Form
Existing Equipment Information	**Required Documentation**
Manufacturer and Model Efficiency: EER, SEER, COP, HSPF Capacity Original fuel type	Photograph and Nameplate Information on Application Form
Quantity	Statement on Application Form
Location/EFLH	Address
Type, quantity, nominal efficiency, and set points of other heating and cooling systems When existing equipment was installed (year) or if this is a new installation	Statement on Application Form
New Equipment Information	**Required Documentation**
Manufacturer and Model Efficiency: SEER and EER (A/C), HSPF (ASHP), or EER and COP (GSHP) Capacity	Cut Sheets
Quantity	Invoice
Location/EFLH	Address
Type, quantity, nominal efficiency, and set points of other heating and cooling systems	Statement on Application Form
For Central Air Conditioners and Air Source Heat Pumps, specify whether the measure is High Efficiency Equipment, Proper Sizing, QIV, Maintenance, or Duct Sealing.	Statement on Application Form
Equipment & Installation Costs	

Inspections should be performed on a random sample of participating customers.

E.2.2. New Construction Programs

Measures: Insulation Upgrades, Efficient Windows, Air Sealing, Efficient HVAC Equipment, Duct Sealing, Efficient Lighting, ENERGY STAR Appliances

New homes will be constructed by participating builders and contractors. The contractors should be required by the utilities to collect and record information about the new homes that receive incentives. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Program managers should create a standard format to document the required information. This will help maintain the uniformity and organization of the documentation throughout each project. It should be the responsibility of the program managers to ensure contractors record the necessary information and that it is complete.

Table E-6: Residential New Construction Program Information to Be Collected

Site Information	Required Documentation
Name of customer Address Phone number Utility and account number (electric)	Statements on Application Form
Area (square feet)	Statement on Application Form
Dimensions and construction details	Construction drawings
Insulation thickness and type	Statements on Application Form
HVAC	Inventory indicating quantities, manufacturers, model numbers, and efficiency ratings for the following equipment, if applicable: furnace, boiler, combination water heater, air source heat pump, geothermal heat pump, PTAC, PTHP, central air conditioning, window air conditioner, thermostat, active solar system
Domestic hot water	Statements on Application Form indicating heater type (electric or gas), size (gallons), and tank insulation type and thickness, if applicable
Lighting	Inventory indicating quantities, manufacturers, model numbers, and wattages
Appliances	Inventory indicating quantities, manufacturers, model numbers, and efficiency ratings
Air infiltration	Blower door test results
Equipment & Installation Costs	

Any information that is available on home energy ratings or other performance ratings conducted through the program should also be collected and stored in the EDCs data tracking and reporting system.

Inspections should be performed on a random sample of new homes.

E.2.3. Energy Star Appliance Programs

Measures: Refrigerator, Freezer, Dehumidifier, Room Air Conditioner, Dishwasher, and Clothes Washer

ENERGY STAR appliances will be provided through participating retailers. The retailers should be required by the utilities to maintain invoices which record information about each appliance that receives an incentive through the program. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Retailers' invoices should include the information listed below. It should be the responsibility of the program managers to ensure retailers record the necessary information and that it is complete.

Table E-7: Residential Energy Star Appliance Program Information to Be Collected

Retailer and Equipment Information	Required Documentation
Name Address Phone number Utility providing incentive Manufacturer and Model Is the model an Energy Star rated model? Type (see below) Quantity When existing equipment was installed (year) or if this is a new installation Existing equipment fuel type (if applicable) Equipment & Installation Costs	Invoices/Application

Refrigerator types: Manual defrost, Partial automatic Defrost, Top mount freezer without door ice, Side mount freezer without door ice, Bottom mount freezer without door ice, Top mount freezer with door ice, and Side mount freezer with door ice.

Freezer types: Upright with manual defrost, Upright with automatic defrost, Chest freezer, Compact upright with manual defrost, Compact Upright with automatic defrost, and Compact chest freezer.
Dehumidifier types: 1-25 pints/day, 25-35 pints/day, 35-45 pints/day, 45-54 pints/day, 54-75 pints/day, and 75-185 pints/day.

Room Air Conditioner – nearest location: Allentown, Erie, Harrisburg, Philadelphia, Pittsburgh, Scranton, and Williamsport.

Dishwasher: Gas or Electric hot water heater.

Clothes Washer: Gas or Electric hot water heater.

E.2.4. Refrigerator/Freezer Retirement Programs

Measures: Refrigerator, Freezer, and Room Air Conditioner[77]

Appliance turn-in incentives will be provided through participating vendors. The vendors should be required by the utilities to maintain invoices and collect record information about each appliance that receives an incentive through the program. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Vendors should collect the information listed below. It should be the responsibility of the program managers to ensure retailers record the necessary information and that it is complete.

[77] Room Air Conditioners are in EDC plans, but not in TRM.

Table E-8: Residential Refrigerator/Freezer Retirement Program Information to Be Collected

Customer and Equipment Information	Required Documentation
Name Address Phone number Utility and account number (electric) Is it an Energy Star rated appliance? Quantity Vendor information Disposal site	Statement on Application Form
Make and Model Capacity	Photograph and Nameplate Information on Application Form
Equipment & Installation Costs	

E.2.5. Energy Star Lighting Programs

Measures: High Efficiency Bulbs and Fixtures[78]

High efficiency lighting measures will be distributed one of two ways: (1) distributed for free at sponsored events and (2) sold at participating retailers. The incentives will be covered by the EDC and paid to the vendor directly.

Vendors should collect the information listed below. It should be the responsibility of the program managers to ensure retailers record the necessary information and that it is complete.

Table E-9: Residential Energy Star Lighting Program Information to Be Collected

Retailer and Equipment Information	Required Documentation
Name Address Phone number Utility providing incentive Manufacturer and Model Type (see Categories below) Wattage Quantity Sold/Distributed Equipment & Installation Costs	Invoices

Categories: CFL Bulbs, Torchieres, Indoor Fixtures, Outdoor Fixture, or Ceiling Fan with ENERGY STAR light fixture

[78] Room Air Conditioners are in EDC plans, but not in TRM.

E.2.6. Energy Star Windows Programs

Measure: ENERGY STAR qualified windows

ENERGY STAR qualified window incentives will be provided to vendors or installers. The vendors or installers should be required by the utilities to collect and record information about the windows that receive incentives. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

Program managers should create a standard format to document the required information. This will help maintain the uniformity and organization of the documentation. It should be the responsibility of the program managers to ensure that vendors or installers record the necessary information and that it is complete.

Table E-10: Residential Window Program Information to Be Collected – if Vendor Receives Incentive

Vendor Information	**Required Documentation**
Vendor name Address Phone number	Statement on Application Form
Window Information	
Manufacturer, Model, and Type U-Factor Solar Heat Gain Coefficient (SHGC) Are the windows Energy Star rated? No. of panes in the new windows	Statement on Application Form
Area (square feet)	Statement on Application Form
Date sold	Statement on Application Form
Equipment & Installation Costs	

Table E-11: Residential Window Program Information to Be Collected – if Installer Receives Incentive

Site Information	Required Documentation
Name of customer Address Phone number Utility and account number (electric)	Statement on Application Form
Heating system type (heat pump, electric resistance, other)	Statement on Application Form
Cooling system type	Statement on Application Form
Existing Window Information	
Manufacturer, Model, and Type	Statement on Application Form
Area (square feet)	Statement on Application Form
Retrofit Window Information	
Manufacturer, Model, and Type U-Factor Solar Heat Gain Coefficient (SHGC) Are the new windows Energy Star rated? No. of panes in the new windows	Statement on Application Form
Area (square feet)	Statement on Application Form
Purchase and installation dates	Statement on Application Form
Equipment & Installation Costs	

Inspections should be performed on a sample of installations.

E.2.7. Direct Load Control Programs

Measures: Air Conditioning Cycling and Pool Pump Load Control

Direct Load Control will be implemented by CSPs. The CSP should be required by the utilities to complete documentation recording information about each customer participating in the direct load control program. The information is necessary to provide accountability and to more accurately assess the gross energy and demand savings of the programs.

The CSP's records should include the information listed below. It should be the responsibility of the program managers to ensure the CSP records the necessary information and that it is complete.

Table E-12: Residential Direct Load Control Program Information to Be Collected

Customer and Equipment Information	**Required Documentation**
Name Address Phone number Utility and account number (electric)	Application Form
Equipment Information	
Manufacturer and Model of controlled equipment Efficiency Capacity Nameplate Wattage	Photograph and Nameplate Information on Application Form
Quantity	Application Form
Equipment & Installation Costs	

The M&V plan for residential direct load programs will be accomplished by monitoring residential utility demand meters during events. Program participants will be expected to reduce their demand by the amount specified in the TRM.

Appendix F. Audit Activities Specific to IPMVP Options

There are specific plausible audit activities for each IPMVP Option, that if necessary, the SWE will conduct. These specific activities are described in the following sections.

F.1. Option A: Partially Measured Retrofit Isolation

In Option A, savings are determined by partial field measurement of the energy use of the system to which the measure was applied; separate from the energy use of the rest of the facility. For these cases the measures are likely to be partially deemed, meaning that some, but not all, parameter(s) are stipulated in the TRM. Thus, the measurement activities are designed to capture the energy use and/or the open variables of the TRM savings equations. These measurements may be short-term or continuous depending on the circumstance and typically patterns of use. Measures that may be applicable for Option A M&V protocols include: high efficiency appliances under a basic level of rigor, HVAC equipment under a basic level of rigor, and lighting efficiency improvements under a basic level of rigor.

Because these measurement activities directly impact the calculated savings of the project, the SWE Team will work closely with the EDCs to ensure that site inspections are carefully planned and executed and that site inspectors have the appropriate experience and training. The SWE Team will audit the following steps in each EDC's site inspection process:

- Training of site inspectors so that they can successfully collect the needed site-specific information.
- Drawing of random sample of sites to be inspected.
- Proper handling of non-response/no-shows/etc.
- Drawing of random and statistically significant sample of measures to measure in cases where multiple measures are implemented in the same project.
- Development of database and site inspection forms.
- Proper use of measurement equipment.
- Performing thorough site inspections and entering all needed data into the program evaluation database.

Since this M&V option is typical of measures with partially deemed savings, the SWE Team will audit the following components used to estimate the savings:

- Use of the TRM formula.
- Use deemed, stipulated, and measured variable values.
- Documentation and calculation of uncertainties.

Table F-1: Auditing IPMVP Option A – Partially measured Retrofit Isolation

Option A Activities	Audit Activities
Pre-retrofit Energy Measure	• Verify that the isolation metering reflects the boundary between equipment which the ECM affects and that which it does not affect.
	• Verify measurement of open variables as defined in M&V Plan.
Verification of Installation	• Verify that conditions of project conform to guidelines of program (i.e., verify eligibility).
	• Verify measure existence.
	• Verify quality of installation.
	• Verify the occurrence of re-inspections post-retrofit regarding ECM performance and operating conditions.
Post-retrofit Energy Measure	• Verify that the isolation metering should reflects the boundary between equipment which the ECM affects and that which it does not affect.
	• Verify measurement of open variables as defined in M&V Plan.
	• Verify frequency of measurement adheres to M&V protocols (e.g., intermittent, short-term, or continuous).
On-Site ECM Sampling	• Verify on-site ECM sampling design adheres to protocols.
	• Review on-site ECM sampling calculations.
	• Verify M&V of ECM sample.
Uncertainties/Independent Variables	• Verify documentation of uncertainties.
	• Review documented uncertainties for accuracy or plausibility.
	• Verify documentation independent variable conditions.
	• Review documented independent variable conditions for accuracy or plausibility.
Instrumentation	• Verify equipment calibration according to NIST procedures.
	• Verify latest measurement techniques according to IPMVP are implemented.
	• Verify protocols for data collection errors and lost data are followed.
Implementation	• Verify the timeliness of participant notification.
	• Verify competency of field-staff.
	• Verify all protocols identified in M&V Plan are followed.
Data Tracking and Reporting	• Verify all variables accounted for in tracking system.
	• QC data tracking and reporting system against EM&V Doc.
Perform Engineering Impact Calculations	• Review project impact calculations.
	• Conduct independent project impact calculations.
Extrapolation of Sample Findings to Impact	• Review sample statistics calculations.
	• Conduct independent sample statistics calculations.

F.2. Option B: Retrofit Isolation

Retrofit isolation measurement protocols are designed to determine savings by field measurement of the energy use of the system to which the measure was applied, separate from the energy use of the rest of the facility. With the short-term or continuous measurements taken on-site, engineering calculations are used to determine the project savings. This type of M&V option is typical of custom projects where the savings are not deemed and there are a number of open variables. Thus, the measurement activities are designed to capture the energy use of the measure, or system within which the measure was implemented, and a few of the key open variables. Measures that may be applicable for Option B M&V protocols include: high efficiency appliances under an enhanced level of rigor, lighting controls under an enhanced level of rigor, and water heating under an enhanced level of rigor.

Because these measurement activities directly impact the calculated or simulated savings of the project, the SWE Team will work closely with the EDCs to ensure that site inspections are carefully planned and executed and that site inspectors have the appropriate experience and training. The SWE Team will audit the following steps in each EDC's site inspection process:

- Training of site inspectors so that they can successfully collect the needed site-specific information.
- Drawing of random sample of sites to be inspected.
- Proper handling of non-response/no-shows/etc.
- Drawing of random and statistically significant sample of measures to measure in cases where multiple measures are implemented in the same project.
- Development of database and site inspection forms.
- Proper use of measurement equipment.
- Performing thorough site inspections and entering all needed data into the program evaluation database.

Since this M&V option is typical of measures with partially deemed savings, the SWE Team will audit the following components used to estimate the savings:

- Use of the TRM formula, appropriate engineering calculations or simulation parameters.
- Use deemed, stipulated, and measured variable values where appropriate.
- Documentation and calculation of uncertainties.

Table F-2: Auditing IPMVP Option B – Retrofit Isolation

Option A Activities	Audit Activities
Pre-retrofit Energy Measure	• Verify that the isolation metering reflects the boundary between equipment which the ECM affects and that which it does not affect.
	• Verify measurement of all variables as defined in M&V Plan.
Verification of Installation	• Verify that conditions of project conform to guidelines of program (i.e., verify eligibility).
	• Verify measure existence.
	• Verify quality of installation.
	• Verify the occurrence of re-inspections post-retrofit regarding ECM performance and operating conditions.
Post-retrofit Energy Measure	• Verify that the isolation metering reflects the boundary between equipment which the ECM affects and that which it does not affect.
	• Verify measurement of open variables as defined in M&V Plan.
	• Verify frequency of measurement adheres to M&V protocols (e.g., intermittent, short-term, or continuous).
On-Site ECM Sampling	• Verify on-site ECM sampling design adheres to protocols.
	• Review on-site ECM sampling calculations.
	• Verify M&V of ECM sample.
Uncertainties/Independent Variables	• Verify documentation of uncertainties.
	• Review documented uncertainties for accuracy or plausibility.
	• Verify documentation independent variable conditions.
	• Review documented independent variable conditions for accuracy or plausibility.
Instrumentation	• Verify equipment calibration according to NIST procedures.
	• Verify latest measurement techniques according to IPMVP are implemented.
	• Verify protocols for data collection errors and lost data are followed.
Implementation	• Verify the timeliness of participant notification.
	• Verify competency of field-staff.
	• Verify all protocols identified in M&V Plan are followed.
Data Tracking and Reporting	• Verify all variables accounted for in tracking system.
	• QC data tracking and reporting system against EM&V Doc.
Perform Engineering Impact Calculations	• Review project impact calculations.
	• Conduct independent project impact calculations.
Extrapolation of Sample Findings to Impact	• Review sample statistics calculations.
	• Conduct independent sample statistics calculations.

F.3. Option C: Whole Building

With the whole building M&V option, savings are determined by measuring energy use at the facility level. These values obtained either with short-term or continuous on-site measurement, can then be used in conjunction with a billing analysis regression model to calibrate the savings estimates resulting from program participation. Measures that may be applicable for Option C M&V protocols include: process improvements and building efficiency upgrades.

Because these measurement activities directly impact the calculated or simulated savings of the project, the SWE Team will work closely with the EDCs to ensure that site inspections are carefully planned and executed and that site inspectors have the appropriate experience and training. The SWE Team will audit the following steps in each EDC's site inspection process:

- Training of site inspectors so that they can successfully collect the needed site-specific information.

- Drawing of random sample of sites to be inspected.
- Proper handling of non-response/no-shows/etc.
- Drawing of random and statistically significant sample of measures to measure in cases where multiple measures are implemented in the same project.
- Development of database and site inspection forms.
- Proper use of measurement equipment.
- Performing thorough site inspections and entering all needed data into the program evaluation database.

Since this M&V option is typical of measures with partially deemed savings, the SWE Team will audit the following components used to estimate the savings:

- Use of the TRM formula, appropriate engineering calculations or simulation parameters.
- Use deemed, stipulated, and measured variable values where appropriate.
- Documentation and calculation of uncertainties.

Table F-3: Auditing IPMVP Option C – Whole Building

Option A Activities	Audit Activities
Pre-retrofit Energy Measure	• Verify that the metering or billing data reflects the energy use associated with the building in which the ECM was installed.
	• Verify data collection of all applicable meters and/or sub-meters.
Verification of Installation	• Verify that conditions of project conform to guidelines of program (i.e., verify eligibility).
	• Verify measure existence.
	• Verify quality of installation.
	• Verify the occurrence of re-inspections post-retrofit regarding ECM performance and operating conditions.
Post-retrofit Energy Measure	• Verify that the metering or billing data reflects the energy use associated with the building in which the ECM was installed.
	• Verify data collection of all applicable meters and/or sub-meters.
	• Verify frequency of measurement adheres to M&V protocols (e.g., intermittent, short-term, or continuous).
On-Site ECM Sampling	• Verify on-site ECM sampling design adheres to protocols (i.e., all sub-meters metered as designated).
	• Review on-site ECM meter reading.
Uncertainties/Independent Variables	• Verify documentation of uncertainties.
	• Review documented uncertainties for accuracy or plausibility.
	• Verify documentation independent variable conditions.
	• Review documented independent variable conditions for accuracy or plausibility.
Instrumentation	• Verify protocols for data collection errors and lost data are followed.
Implementation	• Verify the timeliness of participant notification.
	• Verify competency of field-staff.
	• Verify all protocols identified in M&V Plan are followed.
Data Tracking and Reporting	• Verify all variables accounted for in tracking system.
	• QC data tracking and reporting system against EM&V Doc. and/or billing data.
Perform Regression Analysis	• Review base year model.
	• Assess statistical validity of selected model.
Extrapolation of Data to Determine Impact	• Review savings calculations from regression model.
	• Conduct independent sample statistics calculations.

F.4. Option D: Calibrated Simulation

When the calibrated building energy simulation programs are used as part of the M&V activities, savings are determined through simulation of energy use of components or of the whole facility. Simulation routines must be demonstrated to adequately model actual energy performance measure in the facility. The energy use is typically calibrated with hourly or monthly utility billing data and/or end-uses metering. This method is often used for new construction programs and building, HVAC or shell upgrades in commercial and residential programs. Measures that may be applicable for Option D M&V protocols include: commissioning and operation and maintenance improvements under a basic level of rigor, new construction, and HVAC equipment efficiency under an enhanced level of rigor.

Because these simulation models directly impact the estimated savings of the project, the SWE Team will work closely with the EDCs to ensure that models are carefully designed and created so that they accurately reflect the unique characteristics of the project being analyzed. The SWE Team will audit the following steps in each EDC's model development process, which may include but are not limited to:

- Review of model assumptions
- Verify model capabilities and output
- Validate model energy estimates

The calibrated simulation model option also may involve measurement activities that directly impact the calculated or simulated savings of the project. The SWE Team will work closely with the EDCs to ensure that site inspections are carefully planned and executed and that site inspectors have the appropriate experience and training. The SWE Team will audit the following steps in each EDC's site inspection process:

- Training of site inspectors so that they can successfully collect the needed site-specific information.
- Drawing of random sample of sites to be inspected.
- Proper handling of non-response/no-shows/etc.
- Drawing of random and statistically significant sample of measures to measure in cases where multiple measures are implemented in the same project.
- Development of database and site inspection forms.
- Proper use of measurement equipment.
- Performing thorough site inspections and entering all needed data into the program evaluation database.

Since this M&V option is typical of custom and Partially Deemed Measures, the SWE Team will audit the following components used to estimate the savings:

- Use of the TRM formula, appropriate engineering calculations or simulation parameters.
- Use deemed, stipulated, and measured variable values where appropriate.
- Documentation and calculation of uncertainties.

Table F-4: Auditing IPMVP Option D – Calibrated Simulation

Option A Activities	Audit Activities
Pre-retrofit Energy Measure	• Verify that the metering or billing data reflects the energy use associated with the building in which the ECM was installed.
	• Verify data collection of all applicable meters and/or sub-meters.
Verification of Installation	• Verify that conditions of project conform to guidelines of program (i.e., verify eligibility).
	• Verify measure existence.
	• Verify quality of installation.
	• Verify the occurrence of re-inspections post-retrofit regarding ECM performance and operating conditions.
Post-retrofit Energy Measure	• Verify that the metering or billing data reflects the energy use associated with the building in which the ECM was installed.
	• Verify data collection of all applicable meters and/or sub-meters.
	• Verify frequency of measurement adheres to M&V protocols (e.g., intermittent, short-term, or continuous).
On-Site ECM Sampling	• Verify on-site ECM sampling design adheres to protocols (i.e., all sub-meters metered as designated).
	• Review on-site ECM meter reading.
Uncertainties/Independent Variables	• Verify documentation of uncertainties.
	• Review documented uncertainties for accuracy or plausibility.
	• Verify documentation independent variable conditions.
	• Review documented independent variable conditions for accuracy or plausibility.
Instrumentation	• Verify protocols for data collection errors and lost data are followed.
Implementation	• Verify the timeliness of participant notification.
	• Verify competency of field-staff.
	• Verify all protocols identified in M&V Plan are followed.
Data Tracking and Reporting	• Verify all variables accounted for in tracking system.
	• QC data tracking and reporting system against EM&V Doc. and/or billing data.
Perform Simulation Analysis	• Review base year model.
	• Calibrate simulation to data collected.
	• Assess statistical validity of selected model.
Extrapolation of Data to Determine Impact	• Review savings calculations from simulation model.
	• Conduct independent sample statistics calculations.