**STATE OF INDIANA**

**INDIANA UTILITY REGULATORY COMMISSION**

| | | |
|---|---|---|
| **IN THE MATTER OF THE VERIFIED PETITION OF INDIANAPOLIS POWER & LIGHT FOR APPROVAL OF DEMAND SIDE MANAGEMENT (DSM) PLAN, INCLUDING ENERGY EFFICIENCY (EE) PROGRAMS, AND ASSOCIATED ACCOUNTING AND RATEMAKING TREATMENT, INCLUDING TIMELY RECOVERY THROUGH IPL'S EXISTING STANDARD CONTRACT RIDER NO. 22 OF ASSOCIATED COSTS INCLUDING PROGRAM OPERATING COSTS, NET LOST REVENUE, AND FINANCIAL INCENTIVES.** | ) | **CAUSE NO. ________** |

**PETITIONER'S SUBMISSION OF DIRECT TESTIMONY OF ERIK MILLER**

Indianapolis Power & Light Company ("IPL" or "Petitioner"), by counsel, hereby submits the direct testimony and attachment of Erik Miller.

Respectfully submitted,


Teresa Morton Nyhart (Atty. No. 14044-49)
Jeffrey M. Peabody (Atty. No. 28000-53)
Douglas W. Everette (Atty. No. 34316-49)
BARNES & THORNBURG LLP
11 South Meridian Street
Indianapolis, Indiana 46204

| | |
|---|---|
| Nyhart Phone: | (317) 231-7716 |
| Peabody Phone | (317) 231-6464 |
| Everette Phone: | (317) 231-7764 |
| Fax: | (317) 231-7433 |
| Nyhart Email: | tnyhart@btlaw.com |
| Peabody Email: | jpeabody@btlaw.com |
| Everette Email: | deverette@btlaw.com |

Attorneys for Indianapolis Power & Light Company

**CERTIFICATE OF SERVICE**

The undersigned certifies that a copy of the forgoing was served by electronic transmission on the following:

Indiana Office of Utility Consumer Counselor
115 W. Washington Street, Suite 1500 South
Indianapolis, Indiana 46204
infomgt@oucc.in.gov

Dated this 17th day of May, 2017

Jeffrey M. Peabody

Teresa Morton Nyhart (Atty. No. 14044-49)
Jeffrey M. Peabody (Atty. No. 28000-53)
Douglas W. Everette (Atty. No. 34316-49)
BARNES & THORNBURG LLP
11 South Meridian Street
Indianapolis, Indiana 46204

| | |
|---|---|
| Nyhart Phone: | (317) 231-7716 |
| Peabody Phone | (317) 231-6464 |
| Everette Phone: | (317) 231-7764 |
| Fax: | (317) 231-7433 |
| Nyhart Email: | tnyhart@btlaw.com |
| Peabody Email: | jpeabody@btlaw.com |
| Everette Email: | deverette@btlaw.com |

Attorneys for Indianapolis Power & Light Company

**Petitioner's Exhibit 3**

**PRE-FILED VERIFIED DIRECT TESTIMONY**

**OF**

**ERIK MILLER**

**ON BEHALF OF**

**INDIANAPOLIS POWER & LIGHT COMPANY**

**SPONSORING PETITIONER'S ATTACHMENT EM-1**

# PRE-FILED VERIFIED DIRECT TESTIMONY OF ERIK MILLER

## I. Introduction

**Q1. Please state your name, employer and business address.**

A1. My name is Erik Miller. I am employed by Indianapolis Power & Light Company ("IPL" or "Company"), One Monument Circle, Indianapolis, Indiana 46204.

**Q2. What is your position with IPL?**

A2. I am a Senior Research Analyst.

**Q3. Please briefly describe your educational background and business experience.**

A3. I hold a Bachelor's degree from Indiana University's School of Journalism and a Master of Public Affairs degree from Indiana University's School of Public and Environmental Affairs. Prior to coming to IPL, I worked as a Senior Project Manager for the energy efficiency consulting company, CLEAResult from 2012 – 2015 and prior to that as an Energy Efficiency Program Coordinator at Hoosier Energy Rural Electric Cooperative from 2009 - 2012.

**Q4. What are your current duties and responsibilities at IPL?**

A4. My primary responsibility at IPL is customer end use analysis. This includes customer load forecasting and research, planning and evaluation of DSM programs, and supporting Integrated Resource Planning.

**Q5. Have you previously testified before this Commission?**

A5. Yes. I have previously testified before the Commission in Cause No. 44792, which concerned IPL's DSM programs offered in 2017.

**Q6. Are you sponsoring any attachments in this proceeding?**

A6. Yes. I am sponsoring the following attachment:

Petitioner's Attachment EM-1 IPL Evaluation Framework

**Q7. Did you submit any workpapers?**

A7. Yes. I submitted the electronic spreadsheets underlying my analysis.

**Q8. Were these attachments prepared or assembled by you or under your direction and supervision?**

A8. Yes.

**Q9. What is the purpose of your testimony in this proceeding?**

A9. The purpose of my testimony is to (1) present the cost and benefit analysis of the proposed DSM Plan; (2) discuss how the 2018-2020 DSM Plan Energy Efficiency ("EE") goals are reasonably achievable; consistent with IPL's 2016 Integrated Resource Plan ("IRP"); and designed to achieve an optimal balance of energy resources in IPL's service area; (3) discuss the effect of the proposed DSM Plan on electric rates and customer bills; and (4) describe IPL's plan for conducting evaluation, measurement and verification ("EM&V").

**Q10. Are you familiar with the goals and objectives of DSM?**

A10. Yes, I am. In general, utility offered DSM seeks to influence a customer's demand or consumption of energy supplied by IPL in a manner such that the cost of doing so is more economic than satisfying customer needs through supply-side resources.

## II. Cost And Benefit Analysis

**Q11. Did IPL conduct a cost and benefit analysis of the proposed DSM Plan?**

A11. Yes. The cost and benefit analysis was performed using the Participant Cost Test ("PCT"), Utility Cost Test ("UCT"), Ratepayer Impact Measure ("RIM") Test and the Total Resource Cost Test ("TRC"). These tests are defined in 170 I.A.C. 4-7-1 and discussed below.

**Q12. Please briefly describe each of the tests performed.**

A12. <u>PCT</u>: This test measures the difference between the cost incurred by a participant in a demand-side program and the value received by the participant. A participant's cost includes all costs borne by the participant. A participant's value from a DSM program consists of only the direct economic benefit received by the participant. This test looks at the ratio of the customer bill savings plus the program incentive to their incremental cost to participate in the program. When the value is greater than one, the customer will ultimately save money from program participation.

<u>UCT</u>: The UCT is a cost effectiveness test designed to assess the net present value of a utility's revenue requirements associated with the DSM programs. The test looks at the ratio of the present value of the lifetime benefits (avoided costs) from program delivery to the present value of program delivery costs incurred by the utility. Cost effectiveness is achieved when the ratio of benefits to costs is greater than one.

<u>TRC</u>: This cost effectiveness test eliminates the distinction between a participant and nonparticipant by analyzing whether a resource is cost effective based on the total cost and benefit of the program, independent of the precise allocation to a utility, ratepayer,

and participant. The test looks at the ratio of the present value of the lifetime benefits (avoided costs) from program delivery to the present value of program delivery costs incurred by the utility. The TRC is differentiated from the UCT by reflecting the full incremental cost of the measure in the denominator, regardless of whether the incremental cost is paid for by the utility, customer, or another third party. Cost effectiveness is achieved when the ratio of benefits to costs is greater than one.

RIM: This test analyzes how a rate for electricity is altered by implementing a DSM program. This test measures the change in a revenue requirement expressed on a per unit of sale basis. The test looks at the ratio of the present value of the lifetime benefits (avoided costs) from program delivery to the present values of the program delivery costs incurred by the utility plus the lost revenues to the utility. When the value is less than one, the program is considered to have a negative impact on customer rates in the long term. Note that most energy efficiency program do not pass the RIM test as discussed in my testimony below.

**Q13. For what period of time was the cost and benefit analysis performed?**

A13. The analysis was performed on the lifetime measure impacts and costs for DSM programs delivered in the years 2018-2020.

**Q14. How was cost effectiveness evaluated?**

A14. Programs were evaluated using the DSMore model.

**Q15. What is the DSMore model?**

A15. DSMore is a nationally recognized economic analysis tool developed by Integral Analytics that is specifically designed to evaluate the cost effectiveness of implementing energy efficiency and demand response programs. Unlike many other DSM evaluation tools, the model spreads the savings impacts over distributions of hourly energy prices to provide a robust estimate of the value of DSM.

**Q16. Did you perform the cost effectiveness evaluation?**

A16. Yes.

**Q17. What type of program information was used for the DSMore inputs?**

A17. DSMore inputs include direct program costs (internal administration, vendor implementation, customer incentives, EM&V costs and any incremental customer costs), measure energy and demand savings, measure useful life, net-to-gross ratios and participation rates.

**Q18. Are the costs used in the cost and benefit analysis consistent with Section 10?**

A18. Yes. As previously discussed, IPL evaluated the cost effectiveness of the DSM program portfolio using the standard UCT, TRC, RIM and Participant tests. The types of costs included in the cost and benefit analysis are well established and defined in the California Standard Practice Manual ("CPSC") which is relied on throughout the country, including Indiana.

**Q19. Did IPL include lost revenues in the cost and benefit analysis?**

A19. Yes, when appropriate. In accordance with the CPSC, lost revenue is included in the RIM test and not included in the other standard tests.

**Q20. How were the energy and demand savings associated with each of the program measures determined?**

A20. IPL and the program delivery contractors used the Indiana Technical Reference Manual ("IN TRM") version 2.2 or recent program EM&V to calculate the energy and demand savings by measure. For measures that were not addressed in the IN TRM or EM&V, IPL used TRM resources from nearby states or relied on our understanding of rapidly changing measure attributes, such as lighting costs.

**Q21. What type of utility information do you use for the model inputs?**

A21. Model inputs include avoided costs specific to IPL, as well as customer electricity rates, discount rates, and escalation rates. The avoided capacity, avoided energy and avoided transmission and distribution costs as well as the estimated line loss value used in the analysis are from the 2016 IPL IRP filed on November 1, 2016. Please see Section 5.6.4 of IPL's IRP for additional avoided cost information.

**Q22. Is the proposed 2018-2020 DSM Program portfolio cost effective?**

A22. Yes. The cost effectiveness of the proposed DSM Plan program portfolio, the Residential and Business portfolios and the individual programs as indicated by all four conventional tests, are shown below in Table EM-1.

**Table EM-1 IPL's 2018 – 2020 DSM Plan Cost Effectiveness Results**

| RESIDENTIAL | UCT | TRC | RIM | Participant |
|---|---|---|---|---|
| Appliance Recycling | 2.61 | 2.61 | 0.66 | N/A |
| Community Based Lighting | 2.38 | 2.38 | 0.43 | N/A |
| Residential Demand Response | 2.66 | 4.34 | 2.66 | N/A |
| Income Qualified Weatherization | 0.96 | 0.96 | 0.46 | N/A |
| Lighting & Appliances | 4.43 | 3.85 | 0.71 | 9.69 |
| Multifamily | 1.71 | 1.76 | 0.54 | N/A |
| Peer Comparison | 1.50 | 1.50 | 0.52 | N/A |
| School Education | 3.18 | 3.18 | 0.64 | N/A |
| Whole Home | 1.37 | 1.24 | 0.59 | 4.65 |
| Total | **2.02** | **2.16** | **0.75** | **4.93** |
| | | | | |
| **C&I** | **UCT** | **TRC** | **RIM** | **Participant** |
| Business Custom | 3.53 | 1.85 | 0.82 | 2.89 |
| Business Demand Response | 0.40 | 0.40 | 0.40 | N/A |
| Business Prescriptive | 4.69 | 3.12 | 0.84 | 5.08 |
| Small Business Direct Install | 2.20 | 1.71 | 0.68 | 7.82 |
| Total | **3.61** | **2.34** | **0.81** | **4.05** |
| | | | | |
| **Portfolio** | **2.57** | **2.24** | **0.78** | **4.23** |

*Portfolio and Sector totals include Indirect Costs; Residential = $827,500 /program year, C&I = $827,500 /program year

**Q23. Please describe how the cost effectiveness tests were considered in the DSM Plan development.**

A23. Each test provides unique evaluation criteria for program planning, and IPL reviewed the results of all tests while assembling the 2018 – 2020 DSM Plan.

The PCT indicates whether economically rationale customers will adopt the measures offered in a program. A PCT below 1.0 indicates that a customer will spend more money than they will ultimately save from program participation. Generally, these programs are screened out of the portfolio. Note that there is no incremental cost to the customer to participate in a program which has a PCT result marked as not applicable ("N/A").

IPL also identifies programs that pass the RIM Test. This test provides an indicator of both economic efficiency and fairness among customers. Any program passing this test benefits non-participating customers as well as participating customers in the form of lower rates in the long run and should be considered acceptable. IPL understands that most energy efficiency programs do not pass the RIM test due to the loss in energy sales from savings which are recovered through higher utility rates. If a program fails the RIM test it indicates that rates will likely have to increase. However, the RIM test does not indicate whether rates will increase more if the programs are not implemented. Despite failing the RIM test, these programs may still be offered based on evaluation using the other tests.

IPL also identifies programs that pass both the TRC and the UCT tests. The TRC compares the total costs and benefits of a program for all customers. Program participants benefit through lower bills; whereas, non-participants may be burdened by the costs of the program which are assessed through higher rates. A TRC greater than 1.0 indicates that, on average, all customers benefit.

The UCT assesses the benefits and costs from the utility's perspective by comparing the utility benefits to the utility costs (benefits of avoided energy and capacity costs compared to rebates, incentives and administrative costs) – similar to the Present Value Revenue Requirements as traditionally calculated in the IRP.

It should also be noted that certain programs may not pass the standard benefit-cost tests. However, these programs may have other societal benefits or the benefits are difficult to quantify and have been generally accepted as appropriate DSM programs subject to budget restrictions. Specifically, low-income weatherization programs typically do not pass these cost effectiveness tests, but a DSM program offering for low-income customers is included in the proposed DSM Plan to give such customers the opportunity to participate in programs that will help them manage their energy usage and their energy bills.

**Q24. IPL proposes to continue offering the Business Demand Response Program[1] despite not being cost effective in 2018-2020. Please explain.**

A24. The Business Demand Response Program for 2018 -2020 is intended to maintain the existing level of program participation and in doing so, maintain the Company's investment in this demand side resource. Although Demand Response was not selected by the IRP, IPL included the costs to maintain the existing Air Conditioning Load Management ("ACLM") switch population in the IRP analysis.

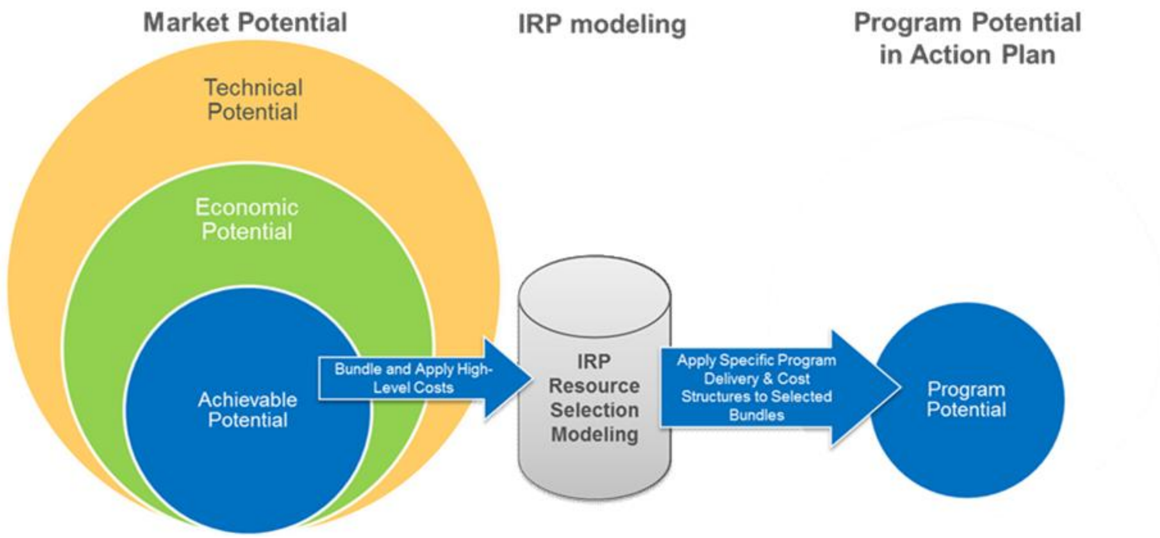
## III. The 2016 IRP and Energy Efficiency Goals

**Q25. Please provide an overview of the planning process IPL has undertaken to arrive at the proposed 2018 – 2020 DSM Plan.**

A25. Existing DSM was reflected as a reduction to the load forecast used in the IRP. Future DSM was modeled using bundles of energy and demand savings of two durations: three years (2018 – 2020) for consistency with the DSM Plan filing period and 16 years (2021 – 2036), the period beyond the filing years. Each bundle was treated as a selectable resource in the IRP model. This process is discussed in Section 5.6 of IPL's 2016 IRP. Figure EM-1 below provides an illustrative representation of IPL's planning process.

[1] Previous to this filing the Business Demand Response program had been identified as the C&I ACLM. The program design is unchanged.

**<u>Figure EM-1</u>**




Note: Representations of Potential Levels are not to scale

**Q26. Please discuss the development of the DSM bundles.**

A26. As further discussed in Section 5.6 of the IRP, IPL began the DSM planning process in early 2016 by completing efficiency DSM Market Potential Study ("MPS") for the 2018 – 2020 period with IPL's consultant Applied Energy Group ("AEG"). After measure-level savings were aligned with recent EM&V or the IN TRM and after IPL's customer load forecast was updated, the MPS estimated different levels of DSM potential in IPL's service territory. These different potential levels included Technical, Economic, Maximum Achievable and Realistic Achievable. IPL used the Maximum Achievable Potential ("MAP") to create bundles of DSM for inclusion in the IRP model. The MAP and RAP give two alternate levels of market adoption rates. Of the two, MAP is the least constrained level of DSM and therefore the appropriate starting point to develop the amount of DSM to be considered in the IRP modeling. Energy efficiency measures in the MAP were then grouped into bundles by similar load shapes and average cost of implementation. The cost tiers were "Up to $30/MWh", "$30-$60/MWh" and "$60+ / MWh". IPL worked with Rick Morgan of the consulting firm Morgan Marketing Partners to use DSMore to create loadshapes and levelized costs for each bundle. The bundles became inputs into the IRP model. Table EM- 2 provides a summary of the 28 DSM bundles that were modeled in the IRP and which bundles were selected by the IRP as a resource for the 2018-2020 period.

**<u>Table EM-2</u>**

| | Levelized Utility Cost per MWh | | |
|---|---|---|---|
| **Sector and Technology** | **(up to $30/MWh)** | **($30-60/MWh)** | **($60+ /MWh)** |
| **EE Residential HVAC** | Selected | Not Selected | Not Selected |
| **EE Residential Lighting** | Selected | N/A | N/A |
| **EE Residential Other** | Selected | Not Selected | Not Selected |
| **EE C&I HVAC** | Selected | Not Selected | Not Selected |
| **EE C&I Lighting** | Selected | Not Selected | Not Selected |
| **EE C&I Other** | Selected | Not Selected | Not Selected |
| **EE C&I Process** | Not Selected | Not Selected | N/A |
| **EE Residential Behavioral** | Not Selected | | |
| **DR Water Heating DLC** | Not Selected | | |
| **DR Smart Thermostats** | Not Selected | | |
| **DR Emerging Tech** | Not Selected | | |
| **DR Curtail Agreements** | Not Selected | | |
| **DR Battery Storage** | Not Selected | | |
| **DR Air Conditioning Load Mgmt** | Not Selected | | |

*N/A indicates that a bundle was not needed; all measures fell within lower cost bundles.

**Q27. What level of DSM was selected in the Company's 2016 IRP?**

A27. As presented in Table EM- 3, the IRP model selected six bundles out of 28 (note: there were a total of 56 bundles available when the bundles for the period 2021 – 2036 are considered). Selected bundles were all in the "Up to $30/MWh" cost tier and included Residential HVAC, Residential Lighting, Residential Other, C&I HVAC, C&I Lighting and C&I Other. The IRP model did not select any Demand Response bundles or incremental new Demand Response; however, IPL included the costs to maintain the existing ACLM switch population in the IRP analysis.

**<u>Table EM- 3</u>**

| Segment | Net Energy Efficiency (MWh) | | |
|---|---|---|---|
| | **2018** | **2019** | **2020** |
| | IRP Selected EE | | |
| Residential | 57,766 | 52,644 | 26,522 |
| Business | 56,638 | 55,073 | 47,664 |
| Total | 114,404 | 107,717 | 74,186 |
| Forecast Sales before incremental utility sponsored EE | 13,769,834 | 13,717,938 | 13,721,071 |

Notes: 2018-2020 estimates were selected in the Resource Selection model in IPL's IRP. Sales forecast is consistent with the load forecast used in IPL's IRP.

**Q28. Did the results of IPL's 2016 IRP modeling prescribe the specific programs to be included in the DSM Plan?**

A28. No. IPL utilizes the IRP to provide long and short-term projections of the optimal balance of supply-side and demand-side resources to reliably meet our customer's energy needs on a reasonable least cost basis. IPL bundled measures by similar load shape and average cost to implement (expressed as $/MWh) rather than bundle by predefined programs. Thus, although IPL's DSM Plan is consistent with IPL's recent IRP, the IRP did not prescribe specific programs be included in the DSM Plan.

The IRP long term projections are at a 20,000 foot level and act as a road map for the utility. As such, the IRP does not specify an absolute known make, model, size and cost for a supply side asset to be built in future planning years. This specific information is determined through an RFP process when the time comes to acquire an asset. Similarly, IPL believes the IRP should not specify exact DSM program designs and strategies because doing so would limit flexibility and the ability to adapt to changing market conditions, costs and technologies. This DSM filing provides the specific program designs, strategies and costs that were developed through the RFP process as discussed by IPL Witness Elliot. As such, the filing reasonably reflects current market information and the passage of time since the MPS was completed.

As discussed by IPL Witness Elliot, in the RFP, IPL provided bidders with parameters around the types of programs to consider in their bids along with savings and costs from the selected IRP bundles. This approach allowed bidders to rely on their expertise in proposing program offerings and customer engagement strategies consistent with IPL's objective to freshen the program offerings.

**Q29. Is the DSM Plan consistent with IPL's most recent long range IRP submitted to the Commission (Section 10(j)(3)(B))?**

A29. Yes. The proposed portfolio in this DSM Plan is designed to be consistent with the IPL 2016 IRP. Table EM-4 below compares the forecasted level of DSM Plan net savings to the level of net savings selected by IPL's 2016 IRP. Overall, the energy savings in the proposed 2018-2020 DSM Plan are slightly greater by 26 GWhs over the three-year planning period, which is consistent with the amount of DSM selected in the IRP process.

**Table EM-4**

| Net Energy Efficiency Savings (GWh) | | | | |
|---|---|---|---|---|
| | 2018 | 2019 | 2020 | Total |
| 2016 IRP | 112 | 106 | 72 | 290 |
| 2018 -2020 DSM Plan | 112 | 112 | 91 | 316 |
| Change | (0) | 7 | 19 | 26 |
| *Savings include Peer Comparison | | | | |

**Q30. Are DSM Plan program operating costs reasonable in light of the cost of DSM selected in the 2016 IRP?**

A30. Table EM-5 below compares the 2018-2020 DSM Plan costs (implementation costs, participant incentive costs and utility administrative costs) to the costs of the DSM selected by the 2016 IRP modeling. For the period 2018–2020, the costs in the DSM Plan proposed herein are approximately $15.8 million more than the Total DSM costs as selected by the IRP.

**Table EM-5**

| Program Delivery Costs (Millions $) | | | | |
|---|---|---|---|---|
| | 2018 | 2019 | 2020 | Total |
| 2016 IRP | $ 18.2 | $ 17.3 | $ 14.9 | $ 50.3 |
| 2018 -2020 DSM Plan | $ 23.6 | $ 23.6 | $ 22.9 | $ 70.1 |
| Change | $ 5.4 | $ 6.3 | $ 8.1 | $ 19.7 |
| *Costs include Peer Comparison and DR | | | | |

**Q31. Please explain why the DSM Plan program operating costs are higher than the DSM costs modeled in the IRP.**

The bids received through the RFP responses are higher than the DSM costs in the most recent MPS. This is expected and normal because the MPS is a theoretical analysis based on the best information available at the time of the analysis. While the costs used in the MPS analysis were conservative and reasonable, the bids received through the RFP process were used to develop a better estimate of the cost to achieve the level of DSM selected by the IRP modeling.

In summary, the DSM Plan portfolio of programs has been modeled in DSMore and is cost effective (except IQW) according to the UCT and TRC using the same avoided costs modeled in the IRP. Based on the results of IPL's analyses, the proposed portfolio in this DSM Plan is consistent with the 2016 IRP.

## IV. <u>Customer Rate Impacts</u>

**Q32. Did IPL consider the effect, or potential effect, in both the long term and short term of the proposed DSM Plan on the electric rates and bills of customers that participate in EE programs compared to the electric rates and bills of customers that do not participate in EE programs (Section 10 (j)(7))?**

A31. Yes. IPL considered stakeholder perspectives when analyzing the cost effectiveness of the 2018-2020 DSM Plan including those of participating customers and non-participating customers. This type of effect is directionally measured by the RIM test which is also called the "non-participant test." Lost revenues, which are assumed to get spread across all customers, are included as a cost in this test. A score less than one indicates that rates will generally go up for all customers. While typically energy efficiency programs score less than one, this test is limited for measuring DSM because it fails to indicate whether rates (over the long term) will increase more than they otherwise would if programs were not implemented. The UCT provides a better indicator of the long run impact to customers by measuring the utility's revenue requirements from the DSM programs. Finally, the Participant Test measures the bill impact to program participants. A score greater than one indicates that a customer's bills will go down as a result of participating in a program. IPL Witness Aliff calculates the DSM Plan bill impact on the typical residential customer using 1,000 kWh per month.

## V. Evaluation, Measurement and Verification

**Q33. Are you familiar with the methodologies used to evaluate DSM?**

A32. Yes. Currently, I oversee the EM&V of IPL's DSM programs which includes ensuring that the third party Evaluator is compliant with the *IPL Evaluation Framework* (see Petitioner's Attachment EM-1) which defines the appropriate methodologies and protocols for evaluating DSM programs.

**Q34. Does the DSM Plan include independent EM&V?**

A33. Yes. IPL will use the *IPL Evaluation Framework*, which was approved by the IPL OSB on June 24, 2015, as a guiding document for the Scope of Work with our third party EM&V vendor. The IPL evaluation plans are designed to meet or exceed the evaluation elements required by 170 IAC 4-8-4. The *IPL Evaluation Framework* also serves as our "plan to assess implementation and quantify the impact on energy and demand of each energy efficiency program and demand response program" as required by the draft 170 IAC 4-8-4 (Evaluation, Measurement and Verification Plan). IPL's Evaluation Framework is included in Petitioner's Attachment EM-1.

IPL intends to issue an RFP for EM&V of the 2018 - 2020 programs described in this filing in the third or fourth quarter of 2017. IPL will keep the IPL DSM OSB informed and provide the OSB an opportunity for input throughout the RFP process.

**Q35. Are the EM&V procedures aligned with applicable environmental regulations, including federal regulations concerning credits for emission reductions (Section10(j)(4))?**

A34. Not at this time. EM&V on utility DSM/EE programs is typically performed at levels specified by the utility based on current, known, requirements. EM&V standards and protocol regarding federal regulations for emission credit reductions are not known at this time. When those requirements are known, IPL will work with both its independent evaluation vendor and OSB to incorporate the requirements needed to comply with any federal and/or state emissions credit plan.

**Q36. Will IPL consider the results of EM&V in determining lost revenues and shared savings?**

A35. Yes. Prior EM&V work performed on IPL programs and the IN TRM, as informed by EM&V, drive the measure level lost revenue and shared savings forecast reflected in this filing. IPL will true-up lost revenues and shared savings based on the most current EM&V when the final annual EM&V report for each Program Year is filed with the Commission. As also discussed by IPL Witness Aliff, this true-up occurs in a semi-annual filing that is made for Standard Contract Rider No. 22 following the conclusion of the annual EM&V.

**Q37. Does this conclude your verified prepared direct testimony?**

A36. Yes.

**VERIFICATION**

I, Erik K. Miller, Senior Research Analyst for Indianapolis Power & Light Company, affirm under penalties of perjury that the foregoing representations are true and correct to the best of my knowledge, information and belief.

Erik K. Miller

Dated: May 17, 2017

# Indianapolis Power & Light Company ("IPL") Evaluation Framework

**April 2015**

## TABLE OF CONTENTS

**INTRODUCTION** .......... **2**
GUIDANCE ON EVALUATION BUDGETING AND BUDGET MANAGEMENT .......... 4
EVALUATION MANAGEMENT, COORDINATION, COMMUNICATION & PROGRESS TRACKING .......... 5
POLICY ON GROSS AND NET SAVINGS AND APPLICATION OF RESULTS .......... 5
BENEFIT COST TESTS AND INPUT METRICS TO TESTS .......... 8
CONTENTS OF EVALUATION REPORTS .......... 10
CYBER SECURITY .......... 11
EVALUATION STANDARDS, ETHICS AND EXPERTISE .......... 12
RISK MITIGATION AND RELIABILITY .......... 14
Bias and Precision .......... 14
Guidelines for assigning value to information .......... 14
Mechanisms for achieving rigor .......... 15
Common sampling approaches .......... 15
90/10 Evaluation confidence and level of precision .......... 16
M&V FIELD PROTOCOLS .......... 17
STANDARDS AND APPROACHES FOR SURVEY RESEARCH .......... 19
ENERGY IMPACT BASELINE APPROACHES .......... 22
Prescriptive Measure Baselines .......... 22
Custom Measure Baselines .......... 22
NET ENERGY IMPACT ATTRIBUTION APPROACHES .......... 30
**APPENDICES** .......... **33**
APPENDIX A: AMERICAN EVALUATION ASSOCIATION GUIDING PRINCIPLES .......... 34
APPENDIX B. DATA NEEDED FOR THE EVALUATION .......... 38
APPENDIX C: ESTABLISHING EFFECTIVE USEFUL LIFE VALUES AND REMAINING USEFUL LIFE 43

# Introduction

This document establishes the Indianapolis Power & Light Company (IPL) DSM Evaluation, Measurement, and Verification ("EM&V") Framework ("IPL Framework" or "Framework").

The purpose of this EM&V Framework is to provide a consistent platform from which evaluations can be designed and implemented so that evaluation results are both reliable and comparable across programs, evaluators, and program implementers. In order to accomplish this purpose this Framework is segregated into two chapters. The first chapter is the Evaluation Policy chapter. The Evaluation Policy chapter provides information pertaining to evaluation-related policies that impact when, how and for what reasons evaluations are conducted. The second chapter is an Evaluation Protocol chapter. The Evaluation Protocol chapter provides information specific to how evaluations are to be conducted.

Evaluation Contractors conducting evaluations of DSM programs should design and implement evaluations that reflect the policy needs presented in the Evaluation Policy chapter and implement evaluations that follow the requirements presented in the Evaluation Protocol chapter.

## Evaluation Objectives

The goal of evaluation is to provide information on the effects of the programs implemented and to provide evidence that can be used to justify cost recovery and help guide future programs and service offerings. This will require flexibility in the evaluation approach so that resources are effectively spent to acquire study results that are reliable, comparable across programs, actionable and which can be used to improve the cost effectiveness of the programs.

## Evaluation and Analysis Approach

Evaluations covered under this Framework include program-specific evaluation efforts, including:

**Impact evaluation –** quantifying the verified gross and net energy savings delivered by programs.

**Process evaluation –**assessing the way in which the programs are designed and implemented, the way they interact within the market, the levels of and drivers for participant satisfaction with the operations and offerings, and other investigative areas.

**Market effects evaluation –**assessing the ways in which energy efficiency programs impact the operations of energy service markets such that additional savings above and beyond those achieved through direct program services to participants are documented.

While written specifically to guide the design and implementation of program-specific energy impact or process evaluation as well as market effects evaluations, this Framework can also provide valuable guidance to the way crosscutting studies are designed and implemented. These types of studies can include the following efforts:

- Market potential studies that assess market baselines and future savings that may be expected for different technologies and customer markets over a specified time horizon.
- Analysis of technology or service gaps that can be met by energy efficiency programs
- Analysis of barriers to energy efficiency implementation and development of approaches to overcome those barriers through redesigned programs
- Action Plans that specify energy saving objectives and methods of achieving those objectives.

## Key EM&V Resource Documents

In addition to this document, there are two EM&V resource documents that will provide the technical basis for planning and conducting evaluation efforts:

1. **Indiana Technical Reference Manual (TRM)** –This document provides the deemed savings estimation approaches and calculation algorithms that should be used in the planning process for program measures.
2. **Industry Standard Protocols** – When not specified in this Framework the Evaluation contractors and their subcontractors (if any) should follow industry standard protocols for best evaluation practice allowed within the resources available. Protocols such as the California Evaluation Protocols[1], the Impact Evaluation Framework for Technology Deployment Programs[2], and other similar publications provide additional perspectives and recommendations for conducting program evaluations. In addition, organizations such as the International Energy Program Evaluation Conference (www.IEPEC.org) publish proceedings containing papers, panels, and presentations on evaluation policy, methods, results and applications that are useful for evaluation professionals

[1] California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals, TecMarket Works, April 2006.
[2] Impact Evaluation Framework for Technology Deployment Programs, USDOE, EERE, July 2007.

# Guidance on Evaluation Budgeting and Budget Management

## Managing the Evaluation Budget to Increase Reliability and Reduce Error Risk

The evaluation budget must be managed to provide the most reliable evaluation results with the lowest probability of error. The EM&V plan should consider the following when developing and approving program-level EM&V approaches and budgets:

- The importance of the program's energy saving contribution to the portfolio. Programs that are expected to provide significant savings should be evaluated using more rigorous approaches than initiatives with lower savings expectations.
- Programs that spend larger portions of the portfolio budget should have a level of evaluation rigor that matches the importance of the program's total financial investment. Thus, larger or more complex programs may have evaluation budgets greater than 5%. However; this increased funding should be off-set by those programs that have evaluation budgets which are lower than 5%.
- Measures with higher level of uncertainty are likely to require higher allocation of budgets. Concentrating effort on measures of high uncertainty will reduce the overall portfolio risk.

Sampling approaches, sample-size targets, and confidence limits should provide the highest level of accuracy achievable within the IPL approved budget. Large programs and programs that are important for reaching energy saving targets should have sampling approaches that reflect that importance. Low impact or smaller programs may have lower precision and confidence levels. However, the precision of the evaluation effort at the program level should be set at 90% confidence and 10% precision levels for a program-cycle[3] unless approved for different levels.

[3] Program cycle: the period of time over which a set of programs are approved for implementation and are subject to a 90/10 level independent evaluation assessment.

# Evaluation Management, Coordination, Communication & Progress Tracking

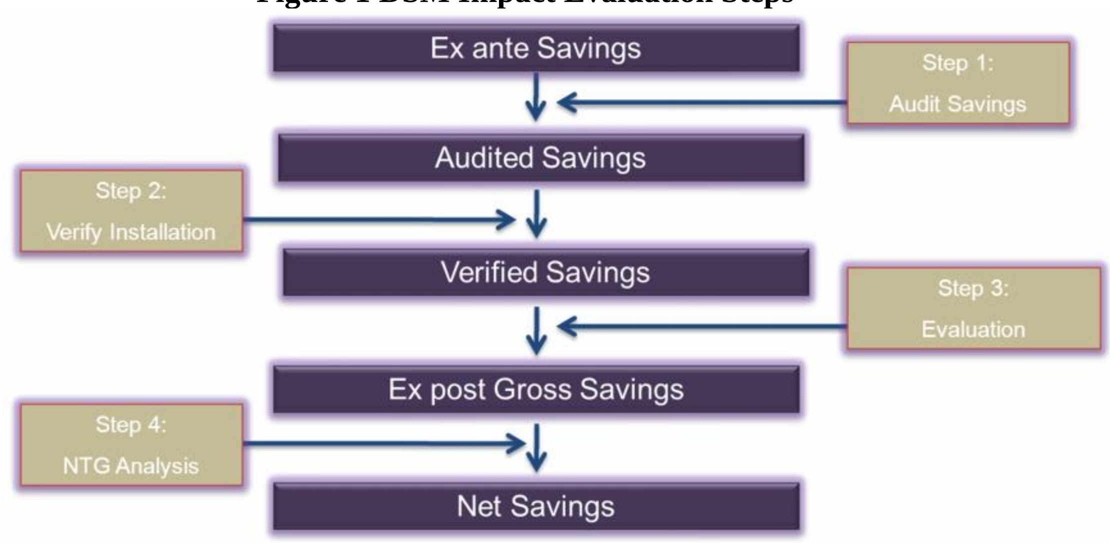
### Progress reporting

It is important that IPL maintains an excellent understanding of the progress and focus of the evaluation activities as they progress. To accomplish this objective the Evaluation contractor should provide periodic progress reports detailing the status and progress of each program evaluation and any crosscutting evaluation effort.

# Policy on Gross and Net Savings and Application of Results

This section describes the typical steps taken in conducting impact evaluations of DSM programs. It also provides definition of different types of energy savings and proposes their appropriate use.

**Figure 1 DSM Impact Evaluation Steps**




### Step 1: Auditing Savings

Validation of the savings claimed within a DSM program will be performed by the evaluation team. The methodology involves the following steps:

1. Reviewing the program tracking databases.
2. Checking saving estimates and calculations against the best available information, (i.e. the adopted Indiana TRM).
3. Reviewing hardcopy program applications from a sample to verify consistency with data recorded in program tracking databases.
4. Adjust program tracking data as necessary to correct any errors, omissions identified in above.
5. Recalculate program savings based on the adjusted program tracking data.

Where custom measures are installed and not part of the TRM, engineering assumptions may be reviewed for a statistically representative sample of projects.
This step results in **Audited Deemed** savings.

### Step 2: Verifying Installations

Step 2 confirms measures have been installed and are operating. This step uses a random sample of installations selected for detailed analysis. Typical methods for collecting necessary data include the following:

1) Telephone Surveys
2) Site Visits

This step may be adjusted to address issues such as:

- Measures rebated but never installed;
- Measures not meeting program qualifications;
- Measures installed but later removed; or
- Measures improperly installed.

Findings from this step produce **Verified Savings.**
*Note: adjustments shown here impact the number of measures reported but do not adjust the TRM saving value.*

### Step 3: Performing Evaluation

At this stage, engineering analysis, building simulation modeling, billing analysis, metering analysis or other accepted statistical methods are used to determine ex post gross savings. Adjustments may include: changes to the baseline assumption; adjustments for weather; adjustments to occupancy levels; adjustments to decreased or increased production levels; and so on. This step does not need to occur annually for every program.
In all cases, the evaluator may use secondary or primary data to perform this step. Secondary data refer to using results from another, similar program, then making minor adjustments for local conditions and installation rates. An example might be using compact fluorescent lamps (CFL) installation rates from a similar utility to adjust the number of bulbs actually installed and saving energy. Secondary data should always be explored as a cost-effective method for adjusting gross savings. Primary data involve collecting information the evaluation requires through surveying program participants, conducting site visits, or metering existing and installed equipment.
*Note: findings reflected from this effort impact the ex post savings reported and may serve as inputs for potential TRM adjustments over time from repetitive ex post studies, but do not adjust the TRM saving value directly.*

### Step 4: Applying NTG

"Net savings" refers to savings directly attributable to a program's efforts. Net savings are determined by adjusting the evaluated gross savings estimates to account for a variety of circumstances, including savings weighted[4] freerider[5] effects, spillover[6] effects and market[7] effects.

The following equations are used to calculate the program's NTG ratio for the two types of net savings estimates:

**Participant Net Savings**

***Annual Net-to-Gross Ratio = (1- freerider adjustment + participant spillover adjustment)***

**Total Net Savings**

***Net-to-Gross Ratio = (1- freerider adjustment + participant spillover adjustment + market effects adjustment)***

For this Framework, three purposes of net savings are identified.

1. To understand the level of net savings achieved by the program and the portfolio to help determine which program to offer in the future.
2. For use in IPL's calculations of lost revenues associated with the energy efficiency and demand response programs.
3. For use in IPL's calculation of Shared Savings incentives associated with the energy efficiency and demand response programs.
4. As a critical evaluation metric to be used for improving program design and implementation. Combined with process evaluations which assess program administration and operations and uncover processes that are ineffective or not well-conceived, the net savings metric assists program implementation toward performance improvements.

Determining the final market effects influenced total net-to-gross (NTG) ratio is not required every year, but should be evaluated every three or four years.

[4] Freerider, spillover and market effects adjustments to the NTG ratio are to be weighted to reflect the level of savings associated with those effects compared to the level of savings that are achieved directly from the installed measures. Savings are weighted so that the adjustments to the net savings are based on the level of savings associated with the actions taken, thus small savings actions result in small adjustments where large savings actions result in larger adjustments, depending on the level of occurrence.
[5] Freeriders are those who would have taken exactly the same action (or made the same behavior change), installing a measure (or changing a behavior) at exactly the same energy efficiency result, at the same time as they took the program-incented action. Partial freeriders are those who would have taken exactly the same action, but the program expedited that change, or they would have taken a similar actions, but not at the same level of efficiency as the program-incented action, or they would have taken the same behavior change but at a later time than the program-encouraged behavior change.
[6] Savings produced as a result of the program's influence on the way participants use energy through technology purchase and use changes or through behavior changes induced or significantly influenced by the program or the portfolio.
[7] Savings produced as a result of the program's or portfolio's influence on the operations of the energy technology markets or changes to energy-related behaviors by customers.

# Benefit Cost Tests and Input Metrics To Tests

## Overview of Benefit-Cost Assessment for DSM Programs

Reference 170 IAC 4-7-7 and 170 IAC 4-8-4. The California Standard Practice Manual (SPM) is the basis for the benefit cost tests in Indiana. In addition, IPL employs an additional test, the Customer Balance Test (CBT) for informational purposes.

## The Total Resource Cost (TRC) Test

The TRC test compares the total costs and benefits of a program for the whole population of customers. The costs include the total costs to the utility and incremental cost of participating customers and the benefits include tax incentives plus the avoided costs of energy supply. The TRC B/C ratio is computed based on the present value of the program benefits (primarily avoided cost of capacity and generation) as well as the total program implementation and operation costs.

Definitions:

• *Incentives.* Incentives are dollar benefits paid by the utility to customers participating in their programs. There are two types of incentives – rebates and rate incentives (monthly bill credits). The rebate type of incentive has the result that the net price to the participant of a program-sponsored device is reduced. A rate incentive is a payment made to reward a participant for his or her behavior. An incentive is defined as being paid directly to the participating customer. As used in this Framework, the term “incentive” includes only rate incentives and direct rebates to customers (which are referred to as “downstream” incentives), and does not include other types of payments that can be made to a variety of entities involved in implementing demand-side programs, such as payments to retailers (referred to as “midstream incentives”) and payments to manufacturers (referred to as “upstream incentives”). Incentives do not include direct install costs of labor or measures.

• *Measure Cost*. Measure cost is the cost of the equipment that is promoted by a particular DSM program. Examples of measure cost include the cost of devices such as energy efficient appliances, switches used to automate a participant’s response to a demand response event, or a solar photovoltaic system. They may also be referred to as *equipment costs*.

• *Incremental measure cost*. Incremental measure cost refers to the difference in cost between a program-sponsored product and an established baseline model of that product (established by codes and standards or by “standard practice”). Energy efficiency cost-effectiveness tests generally use this incremental measure cost, rather than the full equipment cost, because it represents the additional cost that a customer will incur for the energy-efficient product. The incremental measure may be characterized as net of incentives so long as the cost of incentives is included in the utility’s costs as shown in the formula below.

The ratio is usually calculated on a life-cycle basis considering savings and costs that accrue over the lifetime of installed energy efficiency equipment, systems. When the ratio is greater than 1.0, the program is considered cost-effective, with appropriate consideration of uncertainties in the TRC ratio calculation.

$$T \quad T = \frac{A \quad C \quad {}^{*}N \quad E \quad a \quad D \quad S}{P \quad C \quad (i \quad i \quad) + P \quad t's\ i \quad C \quad {}^{8}N \quad o \quad I}$$

## The Utility Cost (UC) Test

The UC test measures the net costs of a program as a resource option based on the costs incurred by the administrator of the program. The benefits are the same as in the TRC test (energy and demand savings value), but the costs are defined more narrowly and do not include consumer costs.

$$U \quad T = \frac{A \quad C \quad {}^{*}N \quad E \quad a \quad D \quad S}{U \quad p \quad C}$$

## The Participant Cost (PCT) Test

The participant test assesses cost effectiveness from the participating consumer's perspective by calculating the quantifiable benefits and costs to the consumer of participating in a program. Since many consumers do not base their decision to participate entirely on quantifiable variables, this test is not necessarily a complete measure of all the benefits and costs a participant perceives.

$$P \quad T = \frac{L \quad R \quad + I}{P \quad C \quad t}$$

## Ratepayer Impact Measure (RIM) Test

The RIM test measures what happens to consumer bills or rates due to changes in utility revenues and operating costs caused by the program. This test indicates the direction and magnitude of the expected impact on rates.

$$R \quad T = \frac{A \quad C \quad {}^{*}N \quad E \quad a \quad D \quad S}{U \quad C \quad + L \quad R}$$

## The Customer Balance (CBT) Test

The CBT is used to assess the degree of subsidization between participants and non-participants. The CBT is not used as a pass/fail test but as a ranking mechanism. Not everyone in the

[8] Note: Participant incremental cost net of incentives is the cost associated with what the participants spent on the energy efficiency project that they would not have spent without the program less the incentives provided by the program. The TRC is to include the participant's cost that are program-induced and not include costs that the participant would have incurred without the program.

customer population receives a net benefit for programs that pass the TRC test. There will be some cross-subsidization between participants and nonparticipants within a customer group but this needs to be minimized to a reasonable extent (see 170 IAC 4-8-5(f)(2)). For example, the TRC ratio can be greater than 1.0 if a small group of participants benefit a great deal at the expense of a large number of non-participants so long as the benefit averaged over all customers is sufficient. This can raise equity issues among customers. For all customers to benefit, the program would need to have a RIM ratio (sometimes called the "no-losers" test) greater than 1.0. This is a difficult standard for most programs. To provide an indication of some balance between these different perspectives, the CBT compares the adverse rate impacts with the aggregate cost savings such that the net benefits of the TRC test must equal or be greater than the net costs of the RIM test. Expressed as a formula:

$$\text{CBT} = \frac{\text{NPV Net Benefits of TRC (Avoided Costs – Utility Costs – Participant Costs)}}{\text{NPV Net Costs of RIM (Utility Costs + Lost revenue – Avoided Costs)}}$$

This ratio, while not eliminating all subsidization between participants and non-participants, does balance the benefits with the total costs which now include rate impacts

# Contents of Evaluation Reports

## Reporting Requirements for Impact, Process, and Market Effects Evaluations

All evaluated gross and net direct energy savings should be reported annually and for the program cycle as a whole, by program, by year. Savings should be reported in three ways, including 1.) ex ante gross, 2.) ex post gross, and 3.) ex post net savings. The reported results should include:

- Electric energy savings kilowatt hours (kWh).
- Electric demand savings (kW).
- Coincident Peak kilowatts (kW).
- When appropriate:
  - Natural gas savings (therms) associated with DSM program measures.
  - And where specifically contracted, therm savings associated with gas measures installed via DSM programs (if any).

Associated with the direct energy savings is the reporting of the following metrics:

- Number of participants and location
- Estimated freerider and spillover percentages (used to calculate net savings)
- Hourly customer usage patterns (obtained for selected programs for which customer on-site metering is conducted)

*Reporting of process evaluation results.* Although the process evaluation efforts will be somewhat different for each program, to a certain extent these studies will follow a similar theme and approach associated with reporting the results of the approved evaluation's scope of effort. That is, the reporting of process evaluation results will depend on the researchable issues on which each evaluation will focus. For this reason we are not identifying the topics on which the evaluation effort will report, however each evaluation report should report the methodological approached used in the process evaluation, the researchable issues on which the evaluation focused, and the findings and recommendations associated with each issue. Findings and recommendations should be numbered so that they can be tracked and referenced and structured to guide program improvement effort. That is, evaluation recommendations should be detailed enough to be well understood and actionable.

Reporting of results may focus on assessment of the following:

- Establishment of the Key Performance Indicators.
- Verification of robust program tracking databases.
- Assessment of participation processes.
- Assessment of market actor interactions/processes.
- Analysis of program design.
- Verification of program processes.

*Reporting of market effects results.* An initial market study will lead to the development of two reports: one on the residential market, and a second for the commercial market. The reports should be cross-cutting by describing the market baseline for multiple end-uses as well as overall market characteristics such as attitudes and barriers towards energy efficiency. Future market effects studies should report changes in the operations of the market and changes to key market change parameters that are caused by the program, and the energy savings associated with those market changes that are program-induced. Energy savings should be reported for the program cycle across the portfolio in the same formats that are required for ex post savings reports. These include:

- Electric energy savings kilowatt hours (kWh).
- Electric demand savings (kW).
- Coincident Peak kilowatts (kW).
- When appropriate:
    - Natural gas savings (therms) associated with DSM program electric measures.
    - And where specifically contracted, therm savings associated with gas measures installed via DSM programs (if any).

# CYBER SECURITY

IPL requires and enforces data security requirements commensurate with the sensitivity of customer identifying information transmitted to EM&V vendors. Data sensitivity and corresponding data transfer and storage requirements are determined at the sole discretion of IPL. All EM&V vendors will be required to complete and submit on a periodic basis, as determined by IPL, a cyber security questionnaire that transparently reports the Evaluator's IT and data security posture.

# Evaluation Standards, Ethics and Expertise

## Evaluation Standards and Ethics

There are a number of evaluation standards and ethics that apply to the evaluation of DSM programs. These standards and ethical considerations guide all evaluation activities covered under this Framework:

**Independence**

The evaluation efforts for DSM programs are to be independent of the DSM program design, approval and service delivery responsibilities. Evaluation contactors can provide support to the DSM program design process by providing evaluation research information, market condition or operations information, program related data, or information needed to support the program design effort.  Evaluation contactors are to maintain an arms-length relationship with the DSM program design, approval and delivery process.

Evaluation efforts are to avoid not only conflicts of interest but also the appearance of conflicts of interests. The evaluators should be independent professionals who do not benefit, or appear to benefit, from the study's findings. The evaluations are also to be independent of program implementers, such that the Evaluation contractor independently develops their study approaches, independently implements those approaches, and independently reports the results from the associated analysis.

**Transparency**

Each evaluation should have a detailed study plan that identifies how the evaluation is to be conducted, specifying the individual tasks within the study to be completed. The study plan should also specify how data will be collected, describe processes to assure objectivity and accuracy, and identify the analysis approach to be applied for each of the four types of evaluation metrics (jobs created, carbon saved, energy demand reduction and energy saved).

The evaluation effort is to be transparent. The methodological description of the study should be sufficiently detailed to allow the research design to be assessed for appropriateness by outside reviewers. The study design should be specific enough to allow other evaluation professionals to understand the approaches used at a sufficient level of detail. The study approach should be transparent to the extent that others can replicate the study approach and obtain similar results. The study plan should also specify how data will be collected, describe processes to assure objectivity and accuracy, and identify the analysis approach to be applied for each of the evaluation objectives.

**Threats to Validity**

The Evaluation contractor should assess the various threats to validity for the study design and analytical approach and develop a study plan that minimizes those threats and reduces the associated level of uncertainty. Both the evaluation plan and the study report should identify these threats and describe how the evaluation approach minimizes any impacts on the study findings.

**Alternative Hypotheses**

To the extent possible, the study design should be developed in a way that addresses alternative hypotheses regarding how observed effects may have occurred.

### Unbiased Assessment

The evaluation design, data collection efforts, analytical approach, and reporting of results should be objective and unbiased. Unsubstantiated claims or unsupported conclusions or personal points of view should be excluded from any evaluation reports or presentations. The study results should be based on objective data/information analysis. Study findings and recommendations should be supported with data and analysis approaches that objectively and impartially assess the available information.

### Attribution of Effects

The study should focus on identifying the outcomes of the projects and programs in question and identify where possible the gross and net effects that can be attributed to the program's efforts.

### Conflict of Interest

Evaluation contractors must disclose any real or perceived conflicts of interest that they might have. These conflicts of interest or perceived conflicts of interest should be identified as a component in the contractor selection process and contractors bidding on the evaluation efforts should present any real or perceived conflicts of interest in their proposals. Likewise, as evaluations evolve and as conditions change within the market, unreported conflicts of interest or potential conflicts of interests should also be brought to the attention of IPL during the course of the evaluation effort as appropriate as they are identified.

A conflict of interest would be reflected in but not necessarily limited to one or more of the following conditions:

1. Any member of the evaluation team or members of their immediate family are a part owner or stockholder or employed by IPL.
2. Any member of the evaluation team or members of their immediate family is employed by an organization who offers energy efficiency program implementation services.
3. Any member of the evaluation team or members of their immediate family is employed by a company or organization owned by or controlled by another organization or company who offers energy efficiency program implementation services.
4. Any member of the evaluation team or members of their immediate family would be in a position to financially benefit from the results of the evaluation findings.

### Sampling

All studies that rely on sampling approaches for collecting data to drive the impact analysis objectives should, to the extent possible, use procedures that minimize bias and maximize the sample's representativeness of the targeted population. Pending the availability of sufficient evaluation budgets, sampling approaches should be structured to be no less rigorous than a 90% level of confidence, per program cycle, with a precision limit of ±10% for the key attributes on which the sample is being selected.

**IPMVP Field Metering and Verification (M&V)[9] Efforts**

Field measurements, when required for assessing equipment baselines and post-retrofit or post installation operations should be conducted using one of the four primary data collection protocols specified in the IPMVP (International Performance Measurement and Verification Protocol). This protocol describes options A, B, C, & D for both single project end use and whole building actions. The IPMVP requires that key performance indicators that drive the estimates of program impacts should be collected via on-site metering, monitoring and verification efforts. The protocol requires measurements to be collected that represent key savings calculation indicators. M&V plans should be developed for each study requiring on-site M&V activities. M&V sampling should be established to be representative of the types of projects and equipment use conditions that represent the largest portion of energy savings. Not all evaluations will require M&V field efforts.

**Survey and Interviews**

When surveys and interviews are used to collect data from which impacts are calculated, the questions should be objective, unbiased and non-leading. Closed-ended, scaled, or quantitative response questions should be structured to allow a full range of applicable responses. Open-ended questions should be single subject response questions that allow for a complete response. Complex questions that require a preamble to set a stage for a response consideration should be avoided to help assure that the response is objective and not guided toward a specific outcome.

# Risk Mitigation and Reliability

## Bias and Precision

**Bias** arises when either the sampling design or the measurement approach leads to estimates that do not equal the true target value (e.g., average savings of population of CFL distributed). In other words, bias is a negative property to be avoided. A confidence interval is a range of values that is believed—with some stated level of confidence—to contain the true population quantity. The confidence level is the probability that the interval actually contains the target quantity. **Precision** provides convenient shorthand for expressing the interval believed to contain the estimator (e.g., if the estimate is 1,000 kWh, and the relative precision level is 10%, then the interval is 1,000 ±100 kWh. Stated another way, we are 90% confident that the true unobserved population value is between 900 and 1,100 kWh).

## Guidelines for assigning value to information

Where resources are limited—*i.e., in nearly every case*—overall validity and precision are optimized by a strategic allocation of effort. Importantly, not all programs need the same level of evaluation rigor. Evaluation budgets should be focused to achieve the most valid and reliable results where they matter most. Evaluation rigor should be matched to the importance of the

[9] M&V refers to Metering and Verification associated with on-site field data collection efforts. The term (M&V) is used differently than the term EM&V in which the E stands for "Evaluation" or the analysis efforts that constitutes the analytical activities within the field of evaluation. Evaluation is the step in which evaluation-related data are analyzed to produce evaluation findings. IMPVP is an M&V effort associated with data collection and operational verification and in itself does not produce evaluation findings but provides the data on which evaluation findings are based.

information being gathered through the evaluation efforts. To achieve this balance the following evaluation rigor considerations are incorporated into the Evaluation Framework:

1. *Contribution to portfolio energy savings*
2. *Share of portfolio budget*
3. *Measure parameter uncertainty*
4. *Expanding programs*
5. *Specific program issues (slow launch, low enrollment, etc.)*
6. *Programs that are known to be ending*
7. *Input from the IPL Oversight Board*

## Mechanisms for achieving rigor

The primary mechanisms by which high levels of rigor are achieved in evaluations include higher sample sizes, frequency of measurement, and estimation methods. Reducing uncertainty usually increases evaluation costs. Thus, research expenditures intended to improve statistical precision should be justified in terms of the value of improved information. Methods of measurement are quite varied but include the metering of equipment on site; on-site inspections without metering; telephone surveys of participants, non-participants, or trade allies; engineering analysis of program data; and review and analysis of secondary data sources. The precision of these methods must be weighed against their relative cost, to achieve an optimal allocation of resources. Likewise, the *number of measurements*, i.e., sample size, and hence the cost, must be balanced against the gains. General principles include:

1. Evaluation planning should focus the type and use of field measurement and verification efforts on those components of the portfolio that have the greatest risk of lowering the precision of the impact estimates.
2. Method selection should consider previous evaluations and the degree of change that has occurred so that as programs change over time, the evaluation focuses additional rigor on programs that have changed.
3. Sampling approaches, sample size targets and confidence limits should be considered so that the effort is focused on improved estimation accuracy or on improving the operations of the programs. For programs that are important components of the efforts should have sampling approaches that reflect that importance.

In addition to the above rigor considerations, at a minimum all statistical precision should match standards outlined in the Indiana TRM. Rigor achieved should also correspond to evaluation reporting criteria.

## Common sampling approaches

The development of the sample requires understanding the necessary accuracy, determining the sample frame, and developing the suitable sampling methodology. Appropriate statistical techniques typically used in energy program evaluation include, but are not limited to:

- ***Simple random sampling:*** drawing randomly from an entire population. This is often, but not always, the most efficient form of sampling.
- ***Stratified sampling:*** drawing randomly from sub-groups within a population. This is used when the variance in a measure is unequally distributed across a population, such as

when the size of savings varies by the size of sites and there is a broad distribution of sizes. Random sampling is done within size groupings.

- ***Ratio sampling:*** sampling to estimate the ratio between two values. This is done, for instance, to estimate a realization rate, where the sample captures both a claimed savings value and a verified savings value. This is not a sampling method, per se, but rather a special use of a sample that affects the sample size. Sampling to estimate a ratio can be more efficient than sampling to estimate a single parameter value.
- ***Nested sampling:*** drawing a sample from within another sample, such as when a site metering sample is drawn from a sample of site verifications.
- ***Systematic sampling:*** often used when a sampling frame is unavailable, such as in store intercept studies. Data is collected at a fixed interval with a random starting point.

### 90/10 Evaluation confidence and level of precision

Energy program evaluation is typically based on estimating energy impacts using a representative sample of program participants to determine how measures are installed and used. The results of these efforts are then used to estimate savings for the program. IPL's DSM program evaluations have a target confidence level of 90% with a relative precision of 10%. How this is applied will depend on several factors, including the need for participant surveys, contractor or trade ally interviews, participant phone verification, on-site verification, on-site metering or monitoring or other data collection approaches for which sampling is constructed. For IPL's evaluations, the evaluation effort should target sampling efforts at key energy estimation metrics to achieve a 90/10 objective. However, a 90/10 objective is not required for all evaluation efforts. The 90/10 standard can be lowered when is not considered beneficial for assessing the researchable issue on which an evaluation objective is based. This provision allows for lower levels of confidence and precision when a 90/10 level is not needed. As a result, a 90/10 objective may be appropriate for assessing the energy impacts of a program, but may not be needed to investigate an objective within the process evaluation. Likewise, a program may be small enough or have a low level of expected savings that the resources used to obtain a 90/10 objective may be better spent increasing the reliability of the findings of a larger program or focusing on a technology with one or more programs that provides larger savings.

This Framework does not specify how the 90/10 objective will be obtained, that is left to the professional discretion of the independent evaluation contactor to determine how best to deploy evaluation resources to achieve the highest level of reliability at the lowest level of estimation error risk at the portfolio level. However the Evaluation contractor should structure their sample at the 90/10 level per program to the extent that this objective can be achieved within the available evaluation budget, and to the extent approved by IPL.

# M&V Field Protocols[10]

This section of the Framework deals with measurement and verification (M&V) protocols, and principles relevant to applying M&V activities for evaluation of DSM programs. Engineering calculations, observation site visits, and metering are techniques that fit together as M&V activities and are used to varying degrees depending on the measure and program and site context. Topics include:

- Overview of M&V
- Selection of an M&V methodology
- Developing the site visit sample
- Quality assurance (QA/QC)
- Training

## Overview of M&V

The following schematic provides an illustrative example of comprehensive M&V.

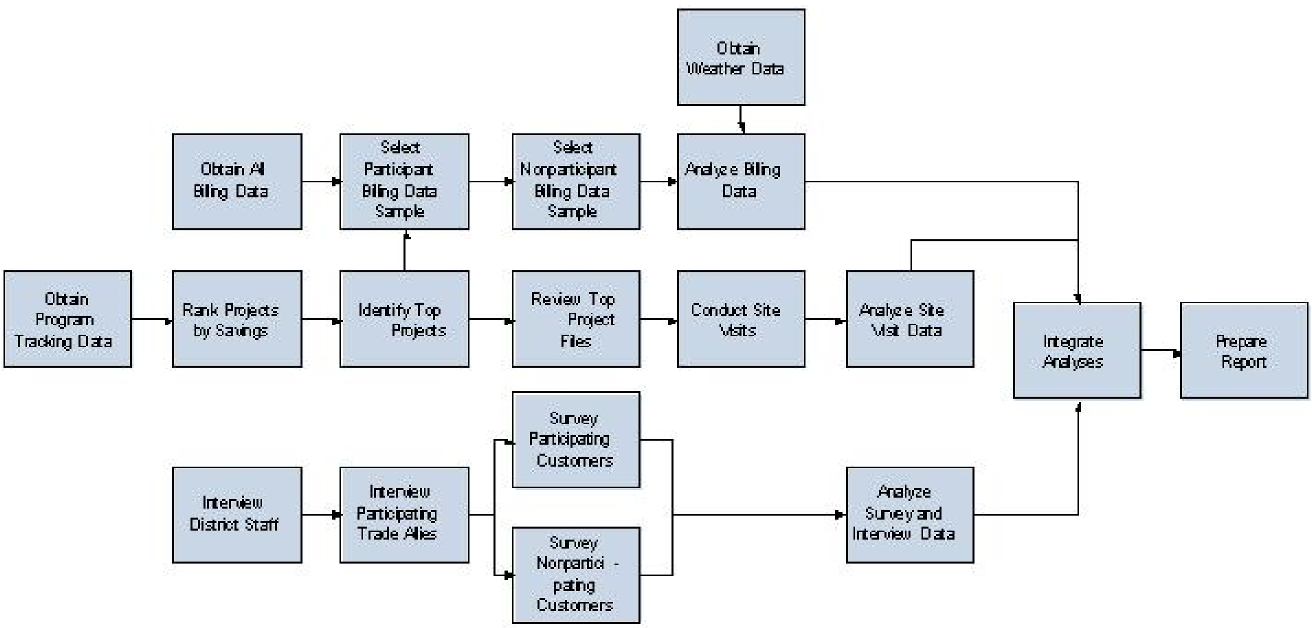



**Figure 2 Comprehensive Monitoring and Verification**

Evaluators generally conduct post-retrofit site visits and associated M&V to determine the savings realization rates associated with a sample of completed DSM projects.

[10] EM&V=Evaluation, Measurement and Verification. EM&V includes the analysis of the collected data (the E component of EM&V). M&V is a limited sub-set of EM&V and is strictly a measurement and equipment operations verification effort.

## Selection of an M&V Methodology

The selection of an M&V methodology or analysis rigor for each sampled site will typically be based on several factors (measure complexity, magnitude of savings, etc.), and this will affect planning for site M&V unit costs accordingly. The following types of on-site verification activities are available to meet the evaluation goals, and will need to be adjusted based on actual site details:

- ***Verification:*** These sites include physical inspection and verification of the operating conditions of the systems under consideration.
- ***Verification with spot measurement:*** These sites involve physical inspection of the installation with spot measurement/reading of the current operating conditions.
- ***Verification with basic rigor:*** These sites will involve meeting–at a minimum–the standards of IPMVP Option A (Partially Measured Retrofit Isolation),[11] including the use of direct measurement.
- ***Verification with enhanced rigor:*** These sites will largely involve using IPMVP Option B (Retrofit Isolation)[12] level analysis.
- ***Phone Survey****:* Call to determine measure presence and operating characteristics.

## Developing the Site Visit Sample

The primary sampling criteria will usually involve stratification of the program population into homogenous groups based on type (e.g., single family vs. multifamily, office vs. retail, etc.), the expected contribution to portfolio savings, and the uncertainty of input variables. Selecting a statistically valid sample is important to the evaluation of DSM programs and requires a complex tradeoff between cost and accuracy.

Evaluators will normally develop the final sampling plan in the first phase of the project and will ensure that the statistical concepts and underlying sampling procedures are clearly explained.

## Quality Assurance and Quality Control

Quality Assurance and quality control (QA/QC) procedures should be set at the inception of the evaluation process: meters should be tested in a metering lab before their use in the field; and nearly all measurements logged should be confirmed using an independent spot- measuring tool—both at installation and at removal—to check logging meter readings. Field staff members should remain on site until all readings are stable. Best practice indicates that all metering points are photographed three times: before the meters are installed, with metering equipment, and after the meters are removed. This allows the evaluation team to confirm equipment nameplates and meter placements after they leave the field.

[11] Savings are determined by field measurement of the key performance parameter(s), which define the energy use of the affected system(s) and/or the success of the project. Measurement frequency ranges from short-term to continuous, depending on the expected variations in the measured parameter and the length of the reporting period.

[12] Savings are determined by field measurement of the energy use of the affected system. Measurement frequency ranges from short-term to continuous, depending on the expected variations in the savings and the length of the reporting period.

# Standards and Approaches for Survey Research

Survey research is a critical piece of the evaluator's toolkit. Nearly all evaluations require the collection or analysis of survey data. This section provides guidance on the design and fielding of structured surveys.

## Principles of Question Wording and Order

A survey is a structured conversation. Like any conversation, word choice can impact understanding. People interpret the same word differently. Survey questions need to be specific, simple and direct; they should address one subject at a time, and need to be exhaustive and mutually exclusive. Questions that will be used in an algorithm to estimate an overall value need to be developed with the algorithm in mind. The algorithm needs to be developed before the survey is designed. The following parts of this section of the Framework provide guidance on survey construction to minimize data bias and improve evaluation reliability.

## Closed-Ended Versus Open-Ended Questions

Surveys typically contain a combination of open- and closed-ended questions. Open-ended questions allow respondents to answer the question in their own words while close-ended questions require respondents to select their response from a provided list.

Close-ended questions are more common because they are easier to administer and analyze and less subject to interviewer effects. Open-ended questions can provide more rich and detailed responses than close-ended questions. However, open-ended questions take longer for respondents answer, require more skilled interviewers, and must be coded for analysis.

A common short-cut is to ask an open-ended question and have the interviewer "field-code" the response by fitting it into pre-defined categories that are not read to the respondent. This approach can reduce analysis time and survey costs, but it is not recommended in most cases. The interviewer becomes the coder and considerable training is typically required for each question to ensure that all interviewers are coding the open-ended responses correctly and consistently. If field-coded open-ended questions are used, long lists of response categories should be avoided as they are difficult for interviewers to manage and can introduce measurement error. Such questions should have no more than five response categories with responses that fall outside these categories typed out in full and recorded as an "other."

Questions that measure a numeric quantity, such as number of CFLs purchased or number of rooms in the house, can and should be asked as an open-ended question. Asking the respondent to fit numeric responses into close-ended category ranges is more likely to produce errors. If ranges are used, the categories should not overlap so that they are mutually exclusive.

## Question Scales

Numeric rating scales are one of the most common question forms. An important decision is the number of scale points. For a scale to provide a reliable and valid measure of a concept, respondents must uniformly understand the meaning of the response categories. Scales with a small number of points are easier for respondents to understand so that respondents tend to interpret the categories in the same manner. The drawback of these scales is that they do not allow finer distinctions in attitudes and behaviors that most respondents are able to make. But scales with too many categories can only provide this higher level of distinction if each point has a clear and distinct meaning. Long scales without clear meaning can create measurement error.

The optimal number of scale points to maximize reliability and validity of survey responses has been the subject of numerous studies. The general consensus is that scales with a moderate number of points – five or seven – tend to have greater reliability and validity than scales with fewer or more points.

### Survey Development and Testing Techniques

Before survey fielding begins, evaluators should employ some form of testing of survey instrument to make sure respondents interpret the questions as intended and are not struggling with the answers.

During the survey development phase, designers could conduct focus groups or cognitive interviews in which the evaluator has the opportunity to talk with respondents to better understand how they interpret the questions. Focus groups and cognitive interviews are time intensive and costly techniques that most are not able to employ. A simple but often overlooked test is to read the survey aloud to someone who was not involved in its development. This exercise will often reveal awkward and confusing wording that can be easily improved.

Once a survey is final and ready for fielding, more formal testing should be conducted. Surveys should be pre-tested with a small number of actual respondents while the evaluator listens to the actual interviews as they are being conducted. Monitoring is one of the only ways a survey designer can hear the full interview from the respondent's perspective. The designer will hear if respondents struggle to understand questions, have difficulty providing an answer that fits the response options, if the interview is too long or repetitive and respondents become impatient compromising data quality.

Evaluators should closely examine the pre-test data to make sure the survey is programmed correctly and respondents are asked all appropriate questions.

All surveys must be reviewed and approved by IPL before fielding begins.

### Survey Fielding

Surveys should be fielded using best practices that are appropriate for the collection mode to ensure minimum bias. For telephone surveys, evaluators should employ call centers that train all new interviewers on proper telephone survey procedures and evaluate the quality of their work on a regular basis. Interviewers should also be trained on the specific survey before they begin calling respondents. The evaluator should explain the purpose of the survey and any unusual or complicated questions.

The survey field period should be long enough so that all sample telephone numbers are dialed numerous times at different times of day to maximize the chance of reaching all respondents. The call center should have procedures for recording the outcome of each call. Ideally, the call dispositions will be recorded in manner that allows the calculation of a response rate using standards set forth by the American Association for Public Opinion Research (AAPOR).

Because mail and Internet surveys are self-administered, evaluators need to pay careful attention to the visual appearance and design of these instruments to minimize respondent error. Evaluators should consider consulting an expert in the field of mail or internet survey design before crafting their field instruments. The field period of mail and Internet surveys should be long enough so that at least one reminder can be sent. The outcome of each email invitation or mailing should also be tracked in a manner to allow the calculation of an AAPOR response rate that is appropriate for internet and mail surveys.

## Survey Methods Reporting

Evaluators should document the survey procedures and methods used so the results can be replicated or compared to other studies. All survey projects should retain:

1. Final survey instruments.
2. A sampling plan that includes a description of the population under study, the sampling frame, the source of the sampling frame, the method used for drawing a sample of respondents from the sampling frame. Any quotas used in fielding the survey should also be detailed.
3. Survey dispositions and response rates. Both should be tracked and calculated using AAPOR Standard Definitions.
4. A description of any survey weights and weight methods.
5. A topline that contains frequency results of all questions asked in the survey.
6. Final data files and computer code used for analysis.

## Ethical Considerations

Evaluators have ethical responsibilities when conducting surveys with utility customers. For each survey, evaluators should inform customers of the sponsor of the survey and that their participation is voluntary. Customers who choose not to answer a question should be respected and not pushed to provide an answer. Any information, alone or in combination, that could identity a customer should be kept confidential unless the customer explicitly waives confidentiality. The Council of American Survey Research Organizations (CASRO) and AAPOR provide codes of standards and ethics. Evaluators must abide by one of these standards. The full CASRO standards can be found at: http://www.casro.org/codeofstandards.cfm. The AAPOR standards can be found at: http://www.aapor.org/AAPOR_Code_of_Ethics/4249.htm.

# Energy Impact Baseline Approaches

## Prescriptive Measure Baselines

The baseline for prescriptive measures should be one of the following:

For early replacement scenario (i.e., replacing existing functioning equipment), the appropriate baseline is the efficiency level of the pre-existing operating equipment. This scenario has another baseline that starts after the end of the remaining useful life (RUL), or when the existing equipment would have ceased to operate. The baseline at that moment is what the customer would have replaced the equipment with, i.e., current market practice or code if the code is enforced. (See Appendix C for detailed discussion of useful lives.)

For non-early replacement scenario (i.e., the equipment is replaced via a new construction program, or for measures where there is no standard RUL identified in this Framework , the baseline is minimum applicable efficiency that is standardly available in the market for that type of equipment or the standard mean market practice or standard mean current practice representing the typical installation. For applications in which there is no building code or appliance standard the baseline is the minimum efficiency level for equipment that is typically installed in similar projects by non-participants. In these conditions the evaluation professional will need to make a judgment call about what is considered minimum efficiency for the range of equipment available in the market. The minimum efficiency equipment (typically called the inefficient choice) represents the lower levels of equipment efficiency available in the market.

**Minimum Efficiency Typically Installed**:
When baseline is set to minimum efficiency, or minimum efficiency level under a code or standards, free rider adjustments are needed to convert gross to net savings. However, it is also possible to set the baseline at a level that includes the influence of freeriders, thus eliminating the need for a freerider adjustment to the gross savings. In this baseline (Standard Market Practice, or SMP) approach, savings are estimated as the difference between the market standard practice baseline and the program induced high efficiency unit. When this approach is used it is assumed that the practice of establishing the market mean practice provides average per measure energy savings that will directly reflect the program's impact net of freeriders. This approach is used when there is a reasonable expectation that participants make decisions similar to those made by non-participants in the absence of the program.

## Custom Measure Baselines

For custom program evaluations the baseline approach can be different for each installation. That is, the technologies as well as the technology configuration and use conditions can be different in each case. As a result, it is not advisable to establish a set of standard baseline approaches. Instead the Framework specifies how project-level baselines can be set, depending on the type of change induced by the program. The evaluation contactor must select the baseline approach appropriate for a set of sampled projects that best reflect the needs of the project and program-level evaluation.

Because there are several different ways that program managers and evaluation experts can define a custom baseline condition, significant differences in savings estimate can result. By defining baselines for various installation conditions, these approaches aim to reduce such differences.

**Types of Custom Projects**
There are typically four types of custom projects.

1. Measures that are not included in the Indiana Technical Reference Manual (TRM) and are unique to a specific non-typical process or application. They are typically not part of prescriptive programs because they do not conform to standard installation and use conditions.
2. Measures not included in the Indiana TRM but are promoted by one or more programs and can be considered a typical installation and therefore should be considered for inclusion in future updates to the Indiana TRM. Because they are not included in the Indiana TRM, custom baseline approaches are needed.
3. Measures that are in the Indiana TRM, but that are installed in a different environment or have a different use conditions than those assumed in the Indiana TRM.
4. Measures that are in the Indiana TRM, but that require simulation modeling or other advanced approaches in order to estimate interactive effects within a facility (if different than category 3 above).

Any one of these four types of custom measures can be mapped into three types (A-C below) of custom projects which require different considerations for estimating pre-program baseline conditions.

A. ***Building*** performance related projects (insulation, space heating, space cooling, domestic water heating, lighting etc.) and,
B. ***Process*** projects that are typically based on the activities that take place within a participant's facilities (paint drying, curing, baking, forming, cutting, stamping, molding, chilling, extruding, compressing, welding, etc.). Space heating and cooling projects are included in the building envelope definitional standard because the performance of these systems is dependent upon both the efficiency and operational conditions of the equipment and conditions of the facility's envelope.

While these two groups work well for many projects, there are also projects that substantially impact post program energy use across both of these groups.

C. ***Building and process*** projects where a change in one significantly impacts the energy use conditions of the other. For example when a facility installs a new high efficiency kiln for drying and forming that is more efficient and better insulated than the previous kiln such that the decreased energy used for baking pottery changes the load on the building's heating and cooling systems. The impacts on the building are the HVAC interactions resulting from the process change.

Within these three types of projects are other considerations for establishing baselines.

**A. Building Projects**

There are two types of building projects: 1) those that are not associated with a building code that is in force at the time of the program-induced change, and 2) those that are covered by a building code which limits the choices that can be considered for the project.

**B. Process Projects**

There are also two types of process projects: 1) those in which the levels of production (i.e., number of units produced annually) increase after installation and 2) those in which they do not increase. Both are further divided into: 1) those not covered by an applicable Federal or state standard, and 2) those covered by an applicable Federal or state standard.

**C. Building and Process Projects**

Some custom projects impact the energy use associated with the operations of the facility and the energy use of certain processes operating within that facility. For these types of projects, baselines must be established for both the facility and the process within the facility. Note that there are cases in which the installation of the installed measure interacts with the energy use of another existing measure (e.g., the installation of a custom lighting measure interacts with the energy use of the existing Heating, Ventilation and Air Conditioning (HVAC) system. In such cases, only the baseline for the installed measure (e.g., lighting) needs to be determined.

**Custom Project Baseline Definitions**

This section defines the baselines for two types of custom *building* projects and four types of custom *process* projects.

1. **Building or facility equipment not covered by a code:** Involves measures associated with the building or facility (envelope, non-deemed and non-process equipment) and measures not covered by a building code. If the program-induced change is an early (before end of life) replacement, the baseline is the pre-program in situ energy consumption. If the program-induced change is a normal replacement (replaced at the end of the effective useful life), the baseline is the energy consumption associated with current practice.
2. **Building or facility equipment that is covered by a code**: Involves measures associated with the building or facility (envelope and non-TRM and non-process equipment) and which are measures covered by a building code that limits the equipment choice. If the program-induced change is an early replacement, the baseline is the pre-program in situ energy consumption. If the program-induced change is a normal replacement, the baseline is the energy consumption associated with current building code.
3. **Process equipment *not covered* by an applicable Federal or state standard**: Involves measures associated with the process or operational activities occurring within the facility that are not covered by an applicable Federal or state standard. If the program-induced change is an early replacement, the baseline is the annual energy consumption of the pre-existing equipment at the post-program level of production. If the program-induced change is a normal replacement, the baseline is the annual energy consumption of equipment representing current practice at the post-installation level of production.
4. **Process equipment *covered* by an applicable Federal or state standard**: Involves measures associated with the process or operations occurring within the facility that are covered by an appliance of equipment standard which limits equipment and change

options. If the program-induced change is an early replacement, baseline is the annual energy consumption of the pre-existing equipment at the post-program level of production. If the program-induced change is a normal replacement, the baseline is the annual energy consumption of equipment that meets the applicable standard at the post-installation level of production.

Note that for numbers three and four above, the issue of whether production increases is irrelevant since the basic assumption is that a given program is not the primary cause of a customer's decision to increase production. There are two reasons supporting this assumption. First, a decision to increase the level of production usually requires a firm to consider a very complex set of organizational and economic factors, only one of which may be the price of electricity and/or gas. Second, to assess whether the program was the primary cause of this decision would require a very complex and prohibitively expensive analysis designed to tease out the effect of the program from the multiple drivers of production changes such as the supply and demand for the firm's product within a national or global market.

In both numbers three and four, the baseline and the post-installation energy use assume the post-installation level of production. This results in greater savings than in the case in which the program is assumed to have caused the increase in the level of production. Both rules recognize that even though the level of production has increased in the post period thereby *increasing* consumption, the *efficiency of production* (kWh/unit) has improved, which has a positive impact on the economic efficiency of the firm and the gross state product.
**Figure 3**~~**Figure 3**~~ below presents the various pathways to defining baselines in each of the types and sub-types discussed above. These definitions also apply to peak kW demand.

**Defining "Current Practice" for Custom Program Baselines**
In determining what constitutes a "*current practice*" in the absence of a building standard or an applicable Federal or state standard, the assessment needs to focus on what equipment choices and installation configurations would have normally been adopted in the absence of the program. (Note: The use of the term current practice should not be confused with the term standard market practice in which a net freerider baseline is defined.) This can be challenging for assessing projects with non-prescriptive measures or for which there is no common per-participant or industry practice which the participant would have followed or that are typical for non-participants. Establishing a *current practice* for a custom project will require some assessment of what each participant would have done in the absence of the program. It is essentially what would have been done without the program assessment. Thus when current per-participant or industry practice is set as the baseline, it is already set at what would have occurred, not as market current practice, but as the custom program participant's current practice. As a result, the impact results are already net of freeriders and no additional freerider adjustment is needed.

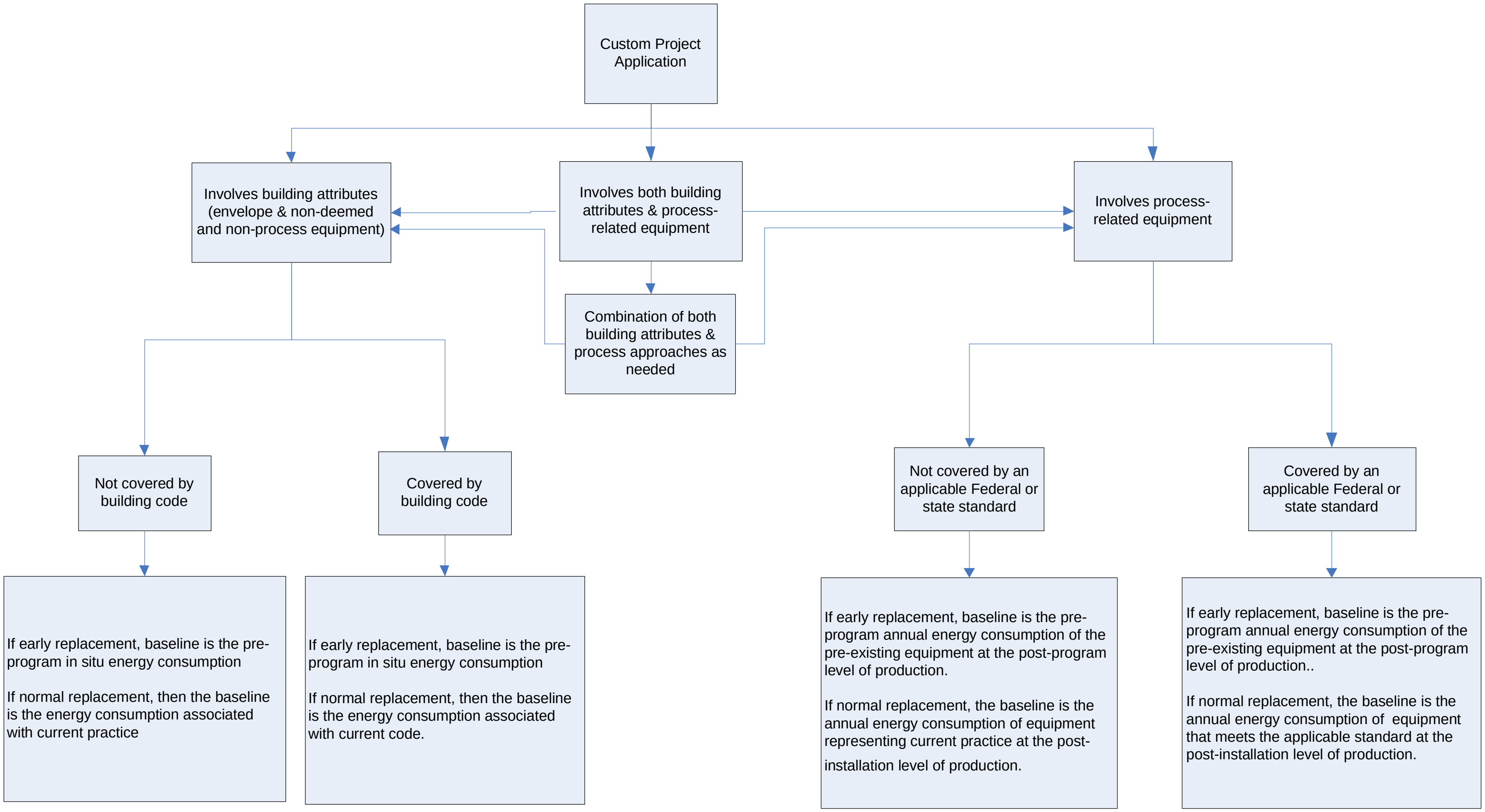
The assessments need to explore a variety of factors affecting what project would have been done in the absence of the program. Factors could include, among other:

- Procurement decision criteria for similar non-program covered equipment;

- The participant's traditional capital investment practices and how they impact equipment choice decisions;
- Past purchase trends for similar equipment;
- Customer self-reports of what they would have installed (if anything) had the program information and incentive not influenced the choice decision;
- Surveys of designers and/or vendors familiar with the process affected by the measure (e.g., interviews with wastewater treatment plant engineers to determine whether variable frequency drivers (VFDs) are common practice on wastewater aerators).

Because energy efficiency programs are designed to influence equipment decisions, one cannot assume that all participants follow what is typically purchased for a specific purpose or use. For many types of custom projects, there may be no typical industry practice. Likewise energy programs are designed to move both early adopters as much as late adopters.

**Figure 3. Determining Baselines for Custom Projects under Various Installation Conditions**



**Effective Useful Life and Remaining Useful Life for Custom Measures**
Since agreed upon effective useful lives (EULs) for general categories of custom projects are not available, case-by-case documentation for the proposed EUL for each custom project should be used in the impact evaluation. Documentation could include dates of installation of the existing equipment that would allow the calculation of its age or, absent such documentation, customer estimates of the age of the existing equipment for each custom project. In some cases, manufacturers' specifications for equipment comprising the custom application could also be used to estimate the EUL. Or, information on time-to-failure of similar equipment supporting similar applications (e.g., plastic extrusion) could be identified within a given industry.

With respect to remaining useful life (RUL), information gathered from knowledgeable people at the site must be gathered to support an estimate of the RUL. For example, such questions as the following could be asked:

- At the time the equipment was replaced, about how many years were left in its useful life (without major repairs which may have led to replacement)?
- Which of the following best describes the condition of the existing equipment when it was replaced: fully functional, fully functioning but with significant problems, or non-functional?
- How long would the old equipment have met the technical and performance needs of the facility?

**Custom Measure Early Replacement:** When a technology is replaced earlier than what would have occurred without the program, the baseline condition is the energy use condition prior to the program-induced change for the remaining useful life of the replaced measure. Once the remaining useful life has expired, the baseline should be established using one of the three methods outlined above and applied to the remaining useful life. In some cases functional application impact calculation adjustments will need to be made by the evaluation contactor when they find that program-caused changes also impact the functions of equipment or processes that are different than the pre-condition.

**Use of Control or Comparison Groups as Baselines**
When the evaluation approach uses experimental or quasi-experimental evaluation approaches[13] the estimation of a pre-program baseline is not required. This is because the participant (test) group's energy use is statistically compared to the consumption of a matched non-participant group (control or comparison group). When random assignment is used to allocate sample points into both the participant and non-participant groups, the difference in consumption between the test and control group provide a net impact result that does not need to be adjusted or modified to provide results that are net of freeriders and participant spillover for that examination period. The same condition applies if quasi-experimental designs are used to establish the test and comparison groups. In both cases the baseline becomes the energy use of the test or comparison group. Experimental designs use random assignments into the two types of groups. Quasi-

[13] Experimental approaches randomly assign people to the participant and control group so that there is theoretically no difference between the two groups. Quasi-experimental approaches build a comparison group (instead of a control group) and statistically control for variable influences that impact the study's findings.

experimental designs use assignments other than random. Quasi-experimental designs are more challenging than experimental design, because differences between the groups that influence energy use need to be controlled statistically.

# Net Energy Impact Attribution Approaches

## Standard Market Practice approach

The standard market practice (SMP) approach is a way to set energy impact analysis baselines so that the baseline already incorporates the influence of freeriders. In this approach a freerider assessment is not needed because the use of a standard market practice baseline is already what the market is doing without the program's direct influence. The SMP baseline is typically set at the mean of the level of energy efficiency being installed across the market being targeted by the program.

## Self-report participant approach

When the SMP approach is not considered to be optimal or appropriate and when experimental or quasi-experimental designs cannot be used, the evaluation should employ a self-reporting approach. The surveys and interview instruments ask a series of questions designed to specifically assess the influence of the program on the participant's decisions. The questions focus on information sources used for making purchase decisions, how the program information influenced the decision, and assessing how the incentive influenced the decision. Participants are also asked about additional actions taken due to the influence of the program, but for which an incentive was not requested or paid. The assessments include consideration for not just the incentives provided, but the information and educational aspects of the program. Net savings can be produced from the incentive, the information provided by a program or the education effects the program has on the purchase and use decision. Each, independently or together, can cause net impacts to be achieved by a program.

The battery of questions used for net analysis are to be kept to a minimum and include only those questions that can reliably be used to estimate net effects. Burdening customers with unnecessary questions that have not been shown to improve the accuracy of an estimation calculation are to be avoided. The development of a standard set of short, focused net-to-gross (NTG) questions will allow the evaluation team to assess freeriders and participant spillover, but will not allow for the addition of market effects.

## Analysis of self-report data

The general analysis approach is to develop an algorithm, based on the direct attribution questions, that establishes an initial attribution factor. Responses to the direct attribution questions will be compared to the context and decision-making questions to identify inconsistencies. The analytical procedures for establishing attribution and for identifying and addressing inconsistencies should be established prior to analysis.

The Evaluation contractor must develop a transparent, straightforward, and readily available matrix approach to assign a score to participants, based on their objective responses to survey questions. Question response patterns are then assigned attribution scores, and the confidence and precision estimates are calculated on the distribution of these scores. The reporting of results should include a matrix (or flow diagram) showing the combinations of responses given to the attribution questions and the percentage of customers (and percentage of the overall savings) that fall into each category. This allows stakeholders to fully understand how each question (and within each question, the response categories) affects the final result. The Evaluation contractor will allow IPL to review the scoring matrix in advance of conducting interviews and share the scoring matrix results once interviews are complete.

The Evaluation contractor's method will also rely on the concept of partial freeridership (partial attribution). Experience has taught evaluation professionals that program participants do not fall neatly into freerider and non-freerider categories. For example, partial freeridership scores were assigned to participants with plans to install the measure; though, the program exerted some influence over their decision, other market characteristics beyond the program also proved influential. In addition, with partial freeridership, we could utilize "Don't Know" and "Refused" responses by classifying them as partial credit, rather than removing the entire respondent from the analysis.  Evaluators then typically weight the respondent freeridership scores by the estimated savings of equipment installed, given the wide variation in nonresidential program participant energy savings.

## Self-report spillover methodology

The concept of spillover refers to additional savings generated by program participants due to their program participation, but not captured by program records. Spillover occurs when participants choose to purchase energy-efficient measures or adopt energy-efficient practices because of a program, but they choose not to participate or are otherwise unable to participate in the program. As these customers are not "participants" for these additional actions, they do not typically appear in program records of the savings generated by spillover impacts. Thus, the energy efficiency programs' spillover effect serves as an additional impact, which can be added to the program's valid results, in contrast to the freeriders' impacts (which reduce net savings attributable to the program).

Evaluations can measure spillover by asking a sample of participants purchasing and receiving a rebate for a particular measure if, due to the program, they installed another efficient measure or undertook other energy efficiency activity. Respondents are typically asked to rate, for example on a scale of 0 through 10, the relative influence of the DSM program and rebate on their decision to pursue additional savings. They may also be asked to explain why they chose not to pursue a rebate for additional measures installed.

Participants are also asked for details regarding the baseline equipment the new energy-efficient equipment replaced. Once the measures and the estimated baseline measures are determined (as best as is feasible within constraints of the survey), detailed measure attributes obtained from the survey questions can be used to establish the most appropriate savings value to assign to that action taken.  In cases where the Indiana TRM do not have applicable energy savings values, the evaluation team will rely on either other accepted values and/or engineering calculations by the evaluation team.

A spillover percentage per program is also calculated by dividing the sum of the additional spillover savings reported by respondents for a given program by total rebated gross savings achieved by all respondents in the program, as follows:

$$Spillover\ \% = \frac{\sum \text{Spillover Measure kWh Savings for All Survey Respondents}}{\sum \text{Program Measure kWh Savings for All Survey Respondents}}$$

## Market effects - non-participant spillover

The evaluations should also assess the level of energy impacts associated with the program's/portfolio's impacts on how the market functions. Energy programs change the way products are selected and priced for sales in areas where energy efficiency programs are operated. These savings are then added to the portfolio's energy savings effects in a way that increases program level savings.

# Appendices

# Appendix A: American Evaluation Association Guiding Principles

A. Systematic Inquiry: *Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.*

1. Evaluators should adhere to the highest appropriate technical standards in conducting their work, whether that work is quantitative or qualitative in nature, so as to increase the accuracy and credibility of the evaluative information they produce.
2. Evaluators should explore with the client the shortcomings and strengths both of the various evaluation questions it might be productive to ask, and the various approaches that might be used for answering those questions.
3. When presenting their work, evaluators should communicate their methods and approaches accurately and in sufficient detail to allow others to understand, interpret, and critique their work. They should make clear the limitations of an evaluation and its results. Evaluators should discuss in a contextually appropriate way those values, assumptions, theories, methods, results, and analyses that significantly affect the interpretation of the evaluative findings. These statements apply to all aspects of the evaluation, from its initial conceptualization to the eventual use of findings.

B. Competence: *Evaluators provide competent performance to stakeholders.*

1. Evaluators should possess (or, here and elsewhere as appropriate, ensure that the evaluation team possesses) the education, abilities, skills, and experience appropriate to undertake the tasks proposed in the evaluation.
2. Evaluators should practice within the limits of their professional training and competence, and should decline to conduct evaluations that fall substantially outside those limits. When declining the commission or request is not feasible or appropriate, evaluators should make clear any significant limitations on the evaluation that might result. Evaluators should make every effort to gain the competence directly or through the assistance of others who possess the required expertise.
3. Evaluators should continually seek to maintain and improve their competencies, in order to provide the highest level of performance in their evaluations. This continuing professional development might include formal coursework and workshops, self-study, evaluations of one's own practice, and working with other evaluators to learn from their skills and expertise.

C. Integrity/Honesty: *Evaluators ensure the honesty and integrity of the entire evaluation process.*

1. Evaluators should negotiate honestly with clients and relevant stakeholders concerning the costs, tasks to be undertaken, limitations of methodology, scope of results likely to be obtained, and uses of data resulting from a specific evaluation. It is primarily the evaluator's responsibility to initiate discussion and clarification of these matters, not the client's.
2. Evaluators should record all changes made in the originally negotiated project plans, and the reasons why the changes were made. If those changes would significantly affect the scope and likely results of the evaluation, the evaluator should inform the client and other

important stakeholders in a timely fashion (barring good reason to the contrary, before proceeding with further work) of the changes and their likely impact.

3. Evaluators should seek to determine, and where appropriate be explicit about, their own, their clients', and other stakeholders' interests concerning the conduct and outcomes of an evaluation (including financial, political, and career interests).
4. Evaluators should disclose any roles or relationships they have concerning whatever is being evaluated that might pose a significant conflict of interest with their role as an evaluator. Any such conflict should be mentioned in reports of the evaluation results.
5. Evaluators should not misrepresent their procedures, data, or findings. Within reasonable limits, they should attempt to prevent or correct any substantial misuses of their work by others.
6. If evaluators determine that certain procedures or activities seem likely to produce misleading evaluative information or conclusions, they have the responsibility to communicate their concerns, and the reasons for them, to the client (the one who funds or requests the evaluation). If discussions with the client do not resolve these concerns, so that a misleading evaluation is then implemented, the evaluator may legitimately decline to conduct the evaluation if that is feasible and appropriate. If not, the evaluator should consult colleagues or relevant stakeholders about other proper ways to proceed (options might include, but are not limited to, discussions at a higher level, a dissenting cover letter or appendix, or refusal to sign the final document).
7. Barring compelling reason to the contrary, evaluators should disclose all sources of financial support for an evaluation, and the source of the request for the evaluation.

D. Respect for People: *Evaluators respect the security, dignity, and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.*

1. Where applicable, evaluators must abide by current professional ethics and standards regarding risks, harms, and burdens that might be engendered to those participating in the evaluation; regarding informed consent for participation in evaluation; and regarding informing participants about the scope and limits of confidentiality. Examples of such standards include federal regulations about protection of human subjects, or the ethical principles of such associations as the American Anthropological Association, the American Educational Research Association, or the American Psychological Association. Although this principle is not intended to extend the applicability of such ethics and standards beyond their current scope, evaluators should abide by them where it is feasible and desirable to do so.
2. Because justified negative or critical conclusions from an evaluation must be explicitly stated, evaluations sometimes produce results that harm client or stakeholder interests. Under this circumstance, evaluators should seek to maximize the benefits and reduce any unnecessary harm that might occur, provided this will not compromise the integrity of the evaluation findings. Evaluators should carefully judge when the benefits from doing the evaluation or in performing certain evaluation procedures should be foregone because of the risks or harms. Where possible, these issues should be anticipated during the negotiation of the evaluation.
3. Knowing that evaluations often will negatively affect the interests of some stakeholders, evaluators should conduct the evaluation and communicate its results in a way that clearly respects the stakeholders' dignity and self-worth.

4. Where feasible, evaluators should attempt to foster the social equity of the evaluation, so that those who give to the evaluation can receive some benefits in return. For example, evaluators should seek to ensure that those who bear the burdens of contributing data and incurring any risks are doing so willingly, and that they have full knowledge of, and maximum feasible opportunity to obtain any benefits that may be produced from the evaluation. When it would not endanger the integrity of the evaluation, respondents or program participants should be informed if and how they can receive services to which they are otherwise entitled without participating in the evaluation.
5. Evaluators have the responsibility to identify and respect differences among participants, such as differences in their culture, religion, gender, disability, age, sexual orientation, and ethnicity, and to be mindful of potential implications of these differences when planning, conducting, analyzing, and reporting their evaluations.

E. <u>Responsibilities for General and Public Welfare:</u> *Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.*

1. When planning and reporting evaluations, evaluators should consider including important perspectives and interests of the full range of stakeholders in the object being evaluated. Evaluators should carefully consider the justification when omitting important value perspectives or the views of important groups.
2. Evaluators should consider not only the immediate operations and outcomes of whatever is being evaluated, but also the broad assumptions, implications, and potential side effects of it.
3. Freedom of information is essential in a democracy. Hence, barring compelling reason to the contrary, evaluators should allow all relevant stakeholders to have access to evaluative information, and should actively disseminate that information to stakeholders if resources allow. If different evaluation results are communicated in forms that are tailored to the interests of different stakeholders, those communications should ensure that each stakeholder group is aware of the existence of the other communications. Communications that are tailored to a given stakeholder should always include all important results that may bear on interests of that stakeholder. In all cases, evaluators should strive to present results as clearly and simply as accuracy allows so that clients and other stakeholders can easily understand the evaluation process and results.
4. Evaluators should maintain a balance between client needs and other needs. Evaluators necessarily have a special relationship with the client who funds or requests the evaluation. By virtue of that relationship, evaluators must strive to meet legitimate client needs whenever it is feasible and appropriate to do so. However, that relationship can also place evaluators in difficult dilemmas when client interests conflict with other interests, or when client interests conflict with the obligation of evaluators for systematic inquiry, competence, integrity, and respect for people. In these cases, evaluators should explicitly identify and discuss the conflicts with the client and relevant stakeholders, resolve them when possible, determine whether continued work on the evaluation is advisable if the conflicts cannot be resolved, and make clear any significant limitations on the evaluation that might result if the conflict is not resolved.
5. Evaluators have obligations that encompass the public interest and good. These obligations are especially important when evaluators are supported by publicly generated funds; but clear threats to the public good should never be ignored in any evaluation. Because the public interest and good are rarely the same as the interests of any particular

group (including those of the client or funding agency), evaluators will usually have to go beyond an analysis of particular stakeholder interests when considering the welfare of society as a whole.

# Appendix B. Data Needed for the Evaluation

This Appendix provides lists of the types of information evaluation contactors may need to support the evaluations of different types of programs.

## Program Information

1. Full program descriptions, including operational or procedures manuals and activities descriptions and description of implementation territories;
2. Detailed descriptions of the tracking system and tracking system operations, including data dictionaries;
3. Program management and staff names, titles, work locations, phone numbers, fax numbers, email addresses;
4. Program theories and associated logic models if developed. If not developed a statement that they have not been developed with a projected date of delivery of the completed theories and logic models;
5. Market operations theories describing the operations of the markets in which the program operates and, if available, a description of how the program is to change the operations of the market;
6. A description of the size of the market targeted by the program, and a description of the baseline conditions at the measure/behavior level and a discussion of how the program is expected to change baseline measure/behavior conditions, if available;
7. A description of the pre-program technical potential at the measure/behavior level and a projection of the remaining technical potential at the end of the program cycle, if available; and
8. When the program relies on key market actors, trade allies and other stakeholders to deliver or support the program in order to reach the energy saving or outreach goals, the TPA should provide a listing, description of and contact information for these individuals/organizations.

## Participant Data

For the purposes of this Framework a participant is defined as an individual or an organization that receives a program service or financial incentive. For most programs, participants are clearly defined in the program tracking systems. However, there are times when a participant is not clearly defined or is not easily identified. Participants signing up for energy efficiency programs are generally easy to identify as they directly receive a service or a financial incentive. Participants in other programs, such as marketing and outreach programs can be harder to identify and report. This Framework does not act to require all programs to identify all participants.

The following participant data should be available in electronic form with supporting database dictionaries to the evaluation teams on request.

## Non-residential program data requests for end-user focused programs

1. Name of program(s) or program component(s);
2. Name of firms participating in program or program component;
3. Service turn on date;
4. Primary and secondary NAICS codes associated with the participants if available;
5. Extent to which customer is a repeat participant or a participant in other programs over the previous five years, if available or accessible;
6. Pre-participation measure and measure-use information, descriptions and conditions;
7. Address(es) of the participating firms or key participation decision makers;
8. Address(es) where program-related action is taken or for the services received;
9. Listing or description of actions taken or services received for each location by measure and end-use according to standard measure and end-use definitions established herein. These lists and descriptions should, to the extent possible, be standardized so that all database developers use the same term for the same measure;
10. Individual participation contact information for each location to include:
    a. First and last name;
    b. Address;
    c. Telephone number;
    d. Fax number (if collected); and
    e. Email address (if collected).
11. Dates of key action/activity/installation steps associated with program participation:
    a. Program enrollment date(s);
    b. Rebate or incentive payment date(s);
    c. Measure install dates;
    d. Date of training received; and
    e. Post-installation measure inspection dates.
12. Financial assistance amounts paid to participant by measure or action taken;
13. Project description information;
14. Estimated savings for actions taken;
15. Summary characteristics of building on which actions are taken or the operational environment in which measures are installed if collected;
16. Account and meter numbers and consumption histories from utility bills from all relevant meters for at least twelve months prior to program enrollment date and through to current period. ;
17. Rate classification; and
18. The size and operational characteristics of the market in which the program is to operate including the number of covered technologies operating in the market and their expected normal failure, change-out or replacement rates.

## Residential program data requests for end-user focused programs

1. Name of program(s) or program component(s) of the participation;
2. Type of building or structure associated with the participant or the participation;
3. Pre-participation measure and measure use information, descriptions and conditions;
4. Service turn on date;
5. Name of individual enrolling in the program or receiving service;
6. Address of the participant;
7. Extent to which customer is a repeat participant or a participant in other programs over the previous five years, if available or accessible;
8. Address where action is taken or for the services received;
9. Listing or description of actions taken or services received according to standard measure and end-use definitions;
10. Individual participation contact information to include:
    a. First and last name;
    b. Address;
    c. Telephone number;
    d. Fax number;(if available and collected); and
    e. Email address (if available and collected).
11. Dates of key action/activity/installation steps associated with program participation:
    a. Program enrollment date(s);
    b. Rebate or incentive payment date(s);
    c. Measure install dates;
    d. Date of training received; and
    e. Post-installation inspection dates.
12. Financial assistance amounts paid to participant by measure or action taken;
13. Project description information;
14. Estimated savings for actions taken;
15. Account numbers and meter numbers and consumption histories from utility bills for all relevant meters for at least twelve months prior to program enrollment date and through to current.
16. Rate classification; and
17. The size and operational characteristics of the market in which the program is to operate including the number of covered technologies operating in the market and their expected normal failure, change-out or replacement rates.

## Non-participant or rejecter data for end-user focused programs

1. Description of program services offered to customer;
2. Date of offering or contact;
3. Method of contact;
4. Name of contact;

5. Address of contact;
6. Telephone number of contact (if known); and
7. Email of contact (if known).

## Program data for mid-stream and upstream focused programs

1. Name of program(s) or program component(s);
2. Name of firms participating in program or program component;
3. Primary and secondary NAICS codes associated with the participants if available;
4. Extent to which customer is a repeat participant or a participant in other programs over the previous five years, if available or accessible;
5. Pre participation/measure and measure use information, descriptions and conditions;
6. Address of the participating firms or key participation decision makers;
7. Address(es) where action is taken or for the services received;
8. Listing or description of actions taken or services received for each location;
9. Individual participation contact information to include:
    a. First and last name (if known) and company name if applicable;
    b. Address;
    c. Telephone number;
    d. FAX number (if collected); and
    e. Email address (if collected).
10. Dates of key action/activity/installation steps associated with program participation:
    a. Program enrollment date(s);
    b. Rebate or incentive payment date(s);
    c. Date of training received; and
    d. Dates, numbers and types of material received.
11. Financial assistance amounts paid to participant by action taken;
12. End-user information as is made available to the program;
13. The size and operational characteristics of the market in which the program is to operate including the number of covered technologies operating in the market and their expected normal failure, change-out or replacement rates; and
14. Names and copies of previous evaluations and market research efforts used by the program to plan and structure program offerings and implementation efforts.

## Program data for information, education and advertising-focused programs

1. Name of program(s) or program component(s);
2. Target population description, size, source of identifying information and lists of population members used in outreach activities. The size and operational characteristics of the market in which the program is to operate including the number of covered technologies operating in the market and their expected normal failure, change-out or replacement rates;
3. Contact information where individual participants are identified to include:

   a. First and last name of key contacts for each location (if known);
   b. Address of individual contacts;
   c. Telephone number of individual contacts;
   d. Fax number of individuals (if collected); and
   e. Email address of individuals (if collected).
4. Marketing materials by numbers, types and distribution;
5. Education or Media plan as appropriate;
6. Execution records for training held; information venues used; program participation agreements, commitments or other similar agreements; post-buy analysis; and other documentation of actual output;
7. Records for dates, number, location, target audience and attendance of events held, Web site hits, call-in numbers and rates, reach, frequency, gross rating points (GRPs), impressions, click through rate, composition, coverage, earned media, value of public service announcements, and other tracking and monitoring information the program maintains, as appropriate to the effort and for each wave, campaign and targeted effort. Include definitions and calculation methods for monitoring statistics used;
8. End-user information available to the program; and
9. Study names and copies of previous evaluations and market research efforts used by the program to plan and structure program offerings and implementation efforts.

# Appendix C: Establishing Effective Useful Life Values and Remaining Useful Life

## Establishing EULs and RULs

The effective useful life (EUL) of an energy efficient measure is the average number of years over which a measure is expected to provide savings. The effective useful life is set is at the estimated point at which 50% of an installed technology type is expected to be remain installed and working in the participant's facilities. Measure lives can vary greatly. An air conditioner installed in a business can last 30 or more years if it is well maintained. In other facilities it may be removed after three years during a remodeling or major equipment up-grade activity. However, it is not uncommon to find measures still installed and performing well beyond their estimated useful life and in some cases for twice the estimated effective useful life. This is because the EUL is set at the *average* number of years the technology is expected to perform.

The remaining useful life (RUL) is the period of time over which the old technology being replaced is expected to have remained in place and functioning if the program would not have been offered to encourage the replacement of that old equipment with a new high efficiency model. The RUL used in evaluation is the expected average RUL across a type or category of technology. In some cases the participant's equipment has failed and is being would have been replaced regardless of the program, in other cases the program can induce a participant to replace the inefficient equipment years before the end of its life.