
The California Evaluation Protocols

A Panel To Summarize And Discuss Their Content And Application

2007 AESP Conference

Nick Hall

Lori Megdal

Pete Jacobs

Rick Ridge

Jane Peters

What the Protocols Do and Why they are Important

- Provides a standardized, uniform approach for documenting program effects.
- Provides increased reliability over previous studies (02-05).
- Sets clear expectations from the evaluation effort.

Where to Find Them

- www.calmac.org
- www.cpuc.ca.gov/static/energyefficiency/rulemaking/eeevaluation/htm
- www.TecMarket.net

Impact Protocol

(Presented by Lori Megdal, Ph.D.)

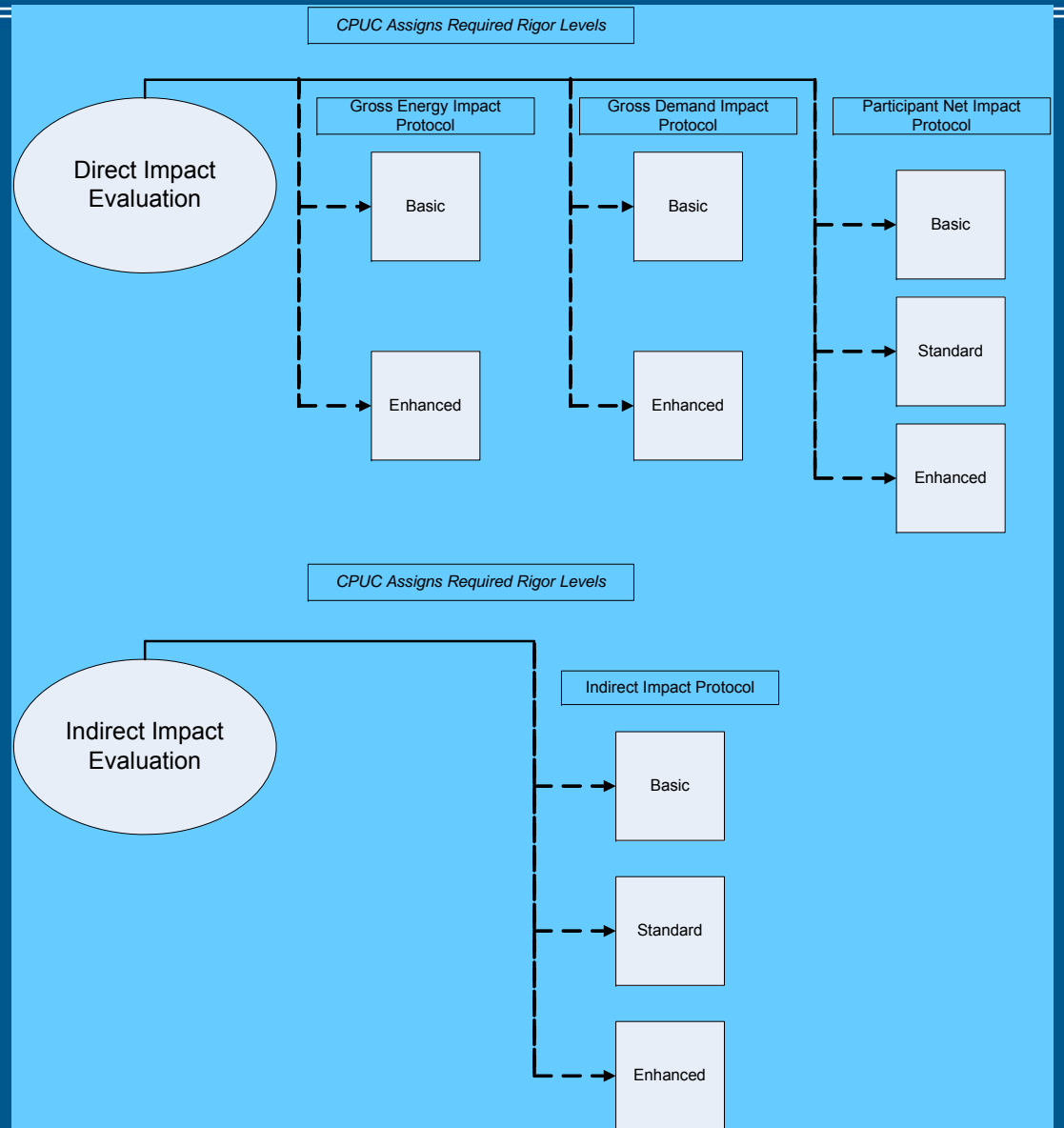
- A primary protocol – for Impact Evaluations
- “Calls” Sampling Protocol
- “Calls” M&V Protocol
- Must meet requirements of Reporting Protocol

Use of Rigor Assignments & Levels of Rigor

- Rigor = Level of expected reliability
- Higher rigor = More confident results are accurate and precise.
- Sampling precision \neq Accuracy
- Rigor level set so minimum standard is reached that includes both precision & accuracy

Impact Evaluation Protocol

Sub-protocols & Rigor Levels



Direct Impact Protocols -- General Differences Due to Rigor Levels

- Gross Energy, Basic: SEM or NAC
- Gross Energy, Enhanced: regression well matched or control for background changes, IPMVP Option D, IPMVP Option B, or Experimental design
- Gross Demand, Basic: End-use load shapes/allocation factors with energy savings
- Gross Demand, Enhanced: Primary data during peak hour peak month for each utility

Participant Net Impact Protocol -- General Differences Due to Rigor Levels #2

- Basic: Participant self-report
- Standard: Participant/non-participant consumption with self-selection adjustment; enhanced self-report with expanded actor, document data; econometric/discrete choice with participant & non-participant addressing self-selection
- Enhanced: “Triangulation” with more than 1 of the methods in the Standard rigor

Indirect Impact Protocol

- Basic Rigor
 - Estimate net changes on participant behavior
- Standard Rigor
 - Estimate net changes on participant behavior
 - Link behavior change to prior study energy & demand savings
- Enhanced
 - Standard Rigor +
 - Field verification of net behavior change

Measurement and Verification

Peter C. Jacobs, P.E.

BuildingMetrics Inc.

Steve Kromer, P.E.

Teton Energy Partners, LLC

M&V Protocol

- M&V is a tool in the overall evaluation process
- “Called” by Impact, Process, ME
- “Calls” Sampling Protocol
- Individual site focus
- Stand-alone guidance for M&V contractors
- Includes measure installation verification
- Basic and enhanced rigor level

Measure Installation Verification Protocol

- Verify correct measure is physically present
- Assess measure installation quality
- Verify correct operation and potential to generate savings (enhanced rigor level only)

M&V Protocol

- Adherence to IPMVP
- Site-specific M&V plan required
- Optimize reduction of uncertainty
- Provide precision estimates
- Relies on and informs DEER where possible

Basic Rigor

- IPMVP Option A – partially measured retrofit isolation
- Stipulated data sources referenced
 - DEER, program workpapers, EM&V literature, catalog data
- Spot or short term measurements as dictated by M&V plan

Enhanced Rigor

- IPMVP Option B (retrofit isolation) or Option D (calibrated simulation)
- Monitoring duration sufficient to capture all operational modes or seasons
- Simulation models calibrated to billing or end-use data
- Calibration criteria apply
- Simulation models compliant with ASHRAE Standard 140-2001

M&V Process

- Obtain site selection and rigor level from impact evaluation plan
- Develop overall M&V plan
- Assess data sources
- Conduct uncertainty analysis
- Develop site-specific M&V plan
- Conduct site monitoring as directed by M&V plan
- Conduct data analysis and estimate site savings
- Conduct uncertainty analysis
- Prepare site M&V report
- Prepare overall M&V report and database

Emerging Technologies

Richard Ridge, Ph.D.
Ridge & Associates

The ETP

- The Statewide Emerging Technologies Program (ETP) is an information-only program that seeks to accelerate the introduction of innovative energy efficient technologies, applications and analytical tools that are not widely adopted in California.
- The overall objective of the ET Program is to verify the performance of new energy efficiency innovations which can be transferred directly into the marketplace and/or integrated into utility portfolios in support of resource acquisition goals for energy efficiency.

Guiding Principles

- Because the ETP and other similar programs will evolve over time, the ETP Protocol is designed to be flexible so that the evaluation, measurement, and verification (EM&V) requirements will apply not only to the 2006-2008 ETP but to future ETP designs as well.
- It is expected that the ETP Protocols will also evolve as evaluators gain experience in evaluating such programs.

Summary of Emerging Technologies

- **2 types of failure– technical performance & utility selection**
- **Provided sample of ETP Evaluation Methods**
- **Required verification of basic achievements**
 - **achieving a certain number of emerging technology application assessments**
 - **updating the Emerging Technology Database, and**
 - **conducting a certain number of meetings annually of the Emerging Technologies Coordinating Council.**
- **The technology application assessments may consist of diverse project types including: feasibility studies, simulation analyses, field demonstrations, controlled environment tests, commercial product development, design methodologies and tool development. Some assessments may take up to four years to complete.**
- **Program theory and logic model**
- **Aggregate level of analysis**
- **Implementation analysis**
- **Analysis of key performance indicators**
- **Peer review**

ETP Reporting Requirements

- Program Theory and Logic Model
- Goal Verification
- Aggregate-Level Analysis
- Implementation Analysis
- Measure Tracking
- Detailed Analysis of Key Performance Indicators
- Peer Review
- Target Audience Surveys

The Market Effects Protocol

Jane Peters

Research Into Action

Market Effects Protocol

- Applies when net market effects are to be estimated at the market level, when:
 - Market effects are used to justify a program
 - Multiple interventions by programs subject to protocols occur in a common market and may influence each other
 - Multiple interventions using other sources occur in a common market and attribution is important to isolate
 - A large program is expected to have broad and long-term market effects in a single market

What the Protocol Does Not Cover

- Free riders – See Impact Protocol
- Participant spillover – see Impact Protocol
- Market progress and assessment – see Process Protocol and market assessment studies
- Estimation of program or measure savings – see Impact and M&V Protocols

What Will Be Reported

- Estimate of the range of net market effects savings in kWh, kW and therms. Also:
 - Market share of energy using equipment or behavior
 - Estimate of non-program induced market share (from naturally occurring and other efforts)
 - Estimate of savings value for net market effect

Key Market Effects Activities

- Scoping study to determine approach and develop evaluation plan
- Baseline and longitudinal indicator measurement
- Analysis of market effects
- Market effects report

Scoping Study

- Define the market
- Market Theory and Logic Model
 - Market operations model integrated with program models
- Identify indicators that need to be tracked
- Identify data sources for both primary and secondary data and how they can be collected
- Develop detailed evaluation plan

Collection of Indicator Data

- Secondary data collection
 - Ensure data choices appropriate to measure indicators of interest
- Primary data collection
 - Experimental or quasi-experimental designs
 - Representative samples of market actors
 - Use of sampling protocol
 - Longitudinal data where possible
 - Sample weighting to represent known population parameters
- Baseline data collected as soon as possible with on-going long-term data collection

Analysis of Market Effects

- Gross market effects
 - Determine change between time one and time two
- Attribution
 - To estimate net effects
 - Preponderance of evidence approach
 - Surveying of samples of knowledgeable market actors from multiple perspectives
 - Net market effects modeling
 - Multivariate analysis of secondary and primary data
- Sustainability
 - Preponderance of evidence approach
- Energy savings
 - Translate market share to savings
 - Direct application of energy savings from impact and M&V studies to market share estimates

Market Effects Evaluation Report

- Document market theory with integrated program theory
- Document data collection activities
- Document estimate of gross market effects
- Document estimate of attribution – net market effects
- Document estimation of net energy market effects

Skills of the Evaluators

- Preponderance of evidence approach
 - Based in qualitative research and survey data collection, analysis uses triangulation
 - Requires significant primary data collection experience, and
 - Significant experience analyzing qualitative data and integration with primary data
- Net Market Effects Modeling
 - Multivariate analysis and regression modeling
 - Likely to use secondary data but some primary data collection probable
 - Requires ability to develop model specification that mimics the market and
 - Understanding of forecasting and other multi-variate analysis approaches

Sampling & Uncertainty

Richard Ridge, Ph.D.
Ridge & Associates

Sampling and Uncertainty Protocol

- Estimation of gross impacts (including M&V);
- Estimation of net impacts;
- M&V in support of specific measure studies; and
- Verification studies in support of non-Impact Evaluation Protocol gross and net impacts.
- Estimation of market effects
- Process evaluations

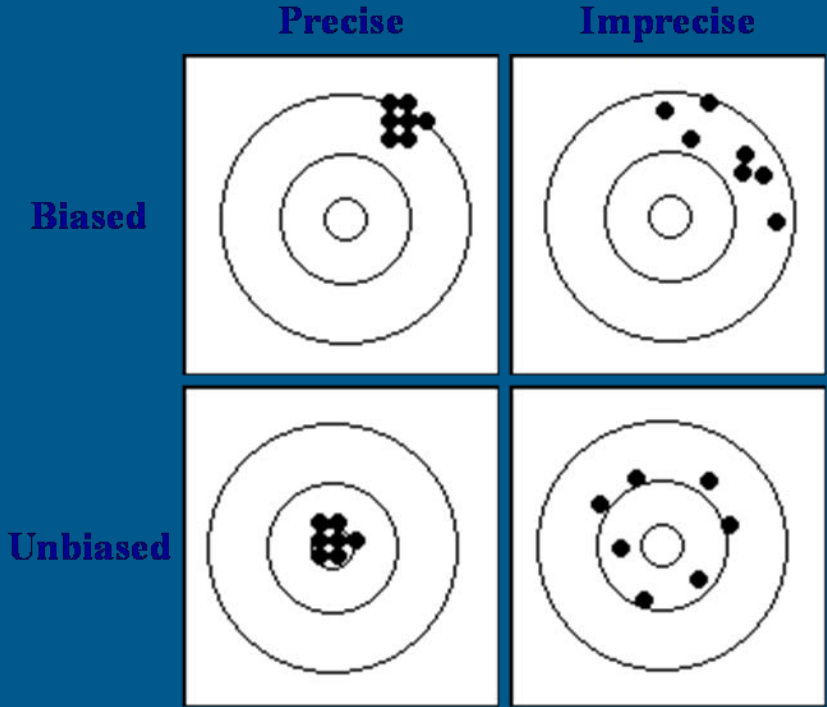
Precision

- Precision is a measure of how close an estimator is expected to be to the true value of a parameter.
- It involves:
 - Degree of confidence
 - Degree of sampling error
 - Degree of variability (Coefficient of Variation (CV), Standard Error, Error Ratio, etc.)
- Examples
 - With a CV of 0.5, 90/10 requires a random sample of 68
 - With a CV of 0.5, 95/5 requires a random sample of 384
- Increasing the CV or the level of confidence increases the required sample size.
- Decreasing the error increases the required sample size.

Bias

- Bias refers to how far the average statistic lies from the parameter it is estimating, i.e., the error which arises when estimating a quantity.
- Types of bias
 - Non-response and other forms of selection bias
 - Measurement bias
 - Erroneous specification of the statistical model.
 - Choosing an appropriate baseline
 - Self selection of program participants
 - Misinterpretation of association as causal effects
 - Construct validity
 - Statistical validity
 - Internal validity
 - External validity

Precision and Bias



Sampling & Uncertainty Principles

- Precision *targets* are set rather than required since,
 - Bias could be much more important than precision for the reliability of the savings estimates or the cost-effectiveness calculations.
 - As any evaluation study proceeds, the data collected could contain much more error than originally thought, requiring more resources to be devoted to reducing this bias and fewer resources devoted to achieving the required statistical precision.
 - Or, the variability in the savings could be so great that it would be impossible to meet the precision requirement.
- The evaluator must have the flexibility to respond to data issues as they arise in order to produce reliable estimates of savings.
- Therefore, focusing on sample error, while giving relatively little attention to these other sources of error, would compromise the CPUC's objective of obtaining *reliable* estimates of kWh and kW impacts.

Precision Overview

- Precision Targets Established for:
 - Gross impact studies
 - Net Impact studies
 - M&V studies
 - Verification studies
 - Process studies
 - Market effects studies
- Depending on the level of rigor and the estimation approaches selection, the precision targets range from 90/10 to 90/30 to none.

Bias Overview

- Cannot ignore such issues as:
 - Non-response and other forms of selection bias
 - Measurement bias
 - Erroneous specification of the statistical model.
 - Choosing an appropriate baseline
 - Self selection of program participants
 - Misinterpretation of association as causal effects
 - Construct validity
 - Statistical validity
 - Internal validity
 - External validity

Process Evaluation Protocol

Nick Hall, TecMarket Works

Process Evaluations

- Largely up to the IOUs to identify need and conduct
- Must present plan to the CPUC
- Must follow the reporting protocol
- CPUC to approve evaluation contractors

The Reporting Protocol

Nick Hall, TecMarket Works

Report Contents

- 1. Cover
- 2. Title Page
- 3. Abstract
- 4. Table of Contents
- 5. Executive Summary
- 6. Introduction and Purpose of the Study.
- 7. Description of Programs Covered in Study
- 8. Study Methodology
- 9. Reliability Assessment
- 10. Detailed Study Findings
- 11. Recommendations for Program Changes
- 12. Appendix A – CPUC performance Metrics
- 13. Appendix B – Data Request Timing & Response

Open Discussion

- Questions
- Comments
- Discussion