Subcommittee Tracking Sheet
Subcommittee Name: Measure Complexity and Best Available information
Meeting #2: February 19, 2015

I. Agenda Items for Discussion/Materials

1. Follow-up from last meeting:
    - Definition of high impact measure
    - 2006-2008 Energy Division Evaluation Plan:
      ""Many evaluator contractors have already begun identifying the measures and/or measure groups within their programs that account for the majority of energy and demand savings from the group of programs that they are evaluating (i.e. "high-impact measures")."
    - 1% of portfolio

2. Relationship between best available information requirements and measure complexity depending on the end use
    - Should quality of data/information inform level of measure complexity? (low impact measures?)
    - Should measure complexity expectations inform whether data/information is sufficient? (high impact measures?)

3. Definition of, and criteria for, "best available information"
    - Settle on how many levels of the standard must be addressed
        - Low impact, medium impact, low impact? (Definition required)
        - How will criteria for different levels differ?
    - Settle on parameters to address in standard for a given source
        - Age
        - Sample size
        - Statistical significance
        - Error band
        - Geographic Origin
        - Evaluability
    - Settle on criteria for application of source to measure methodology
        - Applicability to measure's population
        - High impact versus low impact
        - Conservativism vs optimism
    - Identify additional information needed
        - Common error bands

4. List of sources to be included in "library" of commonly accepted sources

- Settle on list of sources
    1. Other TRMs
    2. E-Source
    3. CEE
    4. Energy Star
    5. Conference papers
- Determine best process for maintaining list
- Caveats that must be included
    1. Always investigate sample size, applicability, etc.

II. Meeting Attendees

TF Members
Doug Mahone
Steven Long

Non-TF Members
Ryan Cho
Mark Gaines
Bhaskar Vempati
Christine Hanhart

Cal TF Staff
Jenny Roecks
Annette Beitel
Alejandra Mejia

III. Key Issues Discussed

a) Prior research that may address statistical standards for robust measure development.

- UN Study: Clean Energy Ministerial: develops standards for reporting data, sponsored by the United Nations Energy Management working group. Provides measurement and verification of energy performance improvements, including metrics for statistical relevance, accuracy, and confidence (7.5% accuracy, 80% confidence).
- Prior CA evaluation: Not a lot of guidance. Evaluation criteria is for ex post evaluation after measure installation, not ex ante forecasts.

Ex ante estimates are used for newer measures for which less information is available.

- UMP: relates to evaluation, not ex ante.
- TRMs
- RTF
- FEMP
- IPMVP

ACT: Review of remaining prior research/initiatives to see if guidelines that would be useful/applicable to ex ante value development.

b) Put more effort into high impact measures.
   - High Impact or high potential measures– Triangulate with multiple methods; use conservative adjustment factors
   - Low Impact – Important to understand ED's expectations for portfolio accuracy (HIMs may account for 20% or 80% of portfolio), will drive accuracy standard for LIMs
   - Greater than 1% of the portfolio is common definition, however don't know right away if or when a new measure will become high impact.
   - Will be helpful to see the impact of PG&E and SCE measures – do the top measures comprise 20% of the portfolio, or 80%? How many "top" measures are there?
   - The NW RTF denotes some measures as "small savers"

   ACT: Look at RTF guidelines for definition and determination of "small saver"
   ACT: Obtain HIM lists from PG&E and SCE examine the portfolio share of top measures.

c) Best Available Information
   - Studies to support WP development tend to be budget-driven and not designed around statistical significance or validity
   - Best available data may not be available until 6-12 mo into program implementation
   - Having WP expiration date can help alleviate concerns about uncertainty in existing data

- A provisional workpaper could specify how data should be collected to refine savings estimate
- Collect data through program implementation, start with an engineering equation or model and calibrate to existing data
- Appropriate level of statistical rigor may differ depending on low or high impact measures

b) What makes measures complex?

- Base case definition: may be a range, individual value, or dual baselines
- Number of building types – multiple lists among utilizes that don't align
- The Energy Commission uses a building type list in its forecast
- Potential for merging CZs for some measures

ACT: Research how NYSERDA addresses dual baselines – they have simplified approach.
ACT: Identify DEER building types; PG&E building types; SCE building types; list of building types that Energy Commission uses in forecast.

c) How to assess the appropriate level of measure complexity

- Current decision rules that don't make sense, create decision tree that makes sense
- Focus on inputs with the most impact
  - Up-front sensitivity analysis
  - Multiple building simulations don't make sense when user behavior is driving factor
  - Draw conclusions from existing measures for similar new measures
- Consider what is a reasonable number of measure combinations to implement successfully base on internal process, customer information
- Consider what drives cost of measures?
  - too many measures to process
  - workpaper development costs (simulations)
  - program budget
- Is the effort to develop simulations worthwhile given the inherent uncertainty in the models and level of achievable accuracy?

- DOE2 models may be 10-15% uncertain, so if outputs of models based on different building types or climate zones differ by less than 10-15%, are multiple measure permutations worthwhile?
- Consideration of effectively communicating to policy makers and regulators
  - Reassurance that utilities are not gaming system
  - Need for savings ranges and not point values in some cases
  - Sensible conservatism
- Need to develop guidelines:
  - Organized critique of current guidelines
  - How do other jurisdictions make these determinations
  - Develop decision tree

IV. Key Action Items

1. Look at RTF guidelines for definition and determination of “small saver”
2. Obtain HIM lists from PG&E and SCE examine the portfolio share of top measures.
3. Research how NYSERDA addresses dual baselines – they have simplified approach.
4. Identify DEER building types; PG&E building types; SCE building types; list of building types that Energy Commission uses in forecast.
5. Review prior research/initiatives to see if guidelines that would be useful/applicable to ex ante value development.
   - UN Study
   - Prior CA evaluation
   - UMP
   - TRMs
   - RTF
   - FEMP
   - IPMVP