Missouri Public Service Commission

POST OFFICE BOX 360
JEFFERSON CITY, MISSOURI 65102
573-751-3234
573-751-1847 (Fax Number)
http://www.psc.mo.gov

May 21, 2013

Richard A. Voytas
Director, Energy Efficiency and Demand Response
Ameren Services
1901 Chouteau Ave - MC 1400
St Louis, MO 63103

Dear Mr. Voytas:

The Staff of the Missouri Public Service Commission first became aware of the planned Cross-Cutting Activities proposed by The Cadmus Group on March 11, 2013. The Activities, and related concerns, were discussed at the March 18-19, 2013 and April 15, 2013 Ameren Missouri stakeholder meetings. While I was not able to make the March meetings, I participated by phone for the April 15 meeting, and met with you, Greg Lovett and John Rogers on May 9, 2013.

Staff has carefully considered the Cross-Cutting Activities evaluation plan, other related information and the resources you provided on May 9. This letter is to inform you that I have discussed this review with Staff and considered the material you presented and concur with John Rogers' representation that Staff cannot support the use of the planned Cross-Cutting Activities proposed by The Cadmus Group for the expressed purpose of adjusting the net-to-gross ratio and annual energy savings of Ameren Missouri's residential energy efficiency programs as part of the evaluation, measurement and verification (EM&V) for the Company's *2013-2015 Energy Efficiency Plan.*

This decision is primarily based on the following facts:

1. Lack of discussion in the Company's *2013 – 2015 Energy Efficiency Plan* of any established industry best practices for estimating energy impacts from non-participants due to market effects as required by Rule 4 CSR 240-22.070(8).
2. Including energy savings for market effects from the Cross-Cutting Study, when that impact was not part of the negotiated Demand-Side Programs Investment Mechanism (DSIM).

Attachment GM-1

Studying market effects is a long-term study which should evolve over several program plan cycles and result in data which is most useful in the development of long-term market potential studies. While Staff does not support the use of the planned Cross-Cutting Activities proposed by The Cadmus Group for the express purpose of adjusting the net-to-gross ratio and annual energy savings through EM&V for program years 2013-2015, Staff does support conducting the Cross-Cutting Activities, and supports the associated budget, in an effort to better understand the market effects specifically due to Ameren Missouri's energy efficiency programs for the purpose of enhancing future planning for demand-side programs and conducting future demand-side market potential studies.

Should you choose to use the planned Cross-Cutting Activities proposed by The Cadmus Group to adjust the net-to-gross ratio and annual energy savings of Ameren Missouri's residential energy efficiency programs as part of the Company's *2013-2015 Energy Efficiency Plan*, Staff reserves its right to challenge that adjustment pursuant to Paragraph 11.b. of the Unanimous Stipulation and Agreement Resolving Ameren Missouri's MEEIA Filing in Case No. EO-2012-0142.

Sincerely,

Natelle Dietrich

Natelle Dietrich
Director – Tariff, Safety, Economic and Engineering Analysis

cc: John Rogers – Missouri Public Service Commission
Greg Lovett – Ameren Missouri

Appendix F

Program-Level Participant and Nonparticipant Spillover Guidance

1. Introduction

The purposes of this Appendix to the *Evaluation Plan Guidance for EEPS Program Administrators* (Guidelines) are to underscore certain methodological principles regarding the reliable estimation of spillover savings, i.e., estimates that are reasonably precise *and* accurate, and to provide additional guidance based on DPS reviews of evaluation plans and completed reports.

The Guidelines define spillover as:

> . . . the energy savings associated with energy efficient equipment installed by consumers who were influenced by an energy efficiency program, but without direct financial or technical assistance from the program. Spillover includes additional actions taken by a program participant as well as actions undertaken by non-participants who have been influenced by the program.

This definition is consistent with the somewhat more detailed definition contained in the California Energy Efficiency Policy Manual (2008):

> Reductions in energy consumption and/or demand in a utility's service area caused by the presence of the DSM program, beyond program related gross or net savings of participants. These effects could result from: (a) additional energy efficiency actions that program participants take outside the program as a result of having participated; (b) changes in the array of energy-using equipment that manufacturers, dealers and contractors offer all customers as a result of program availability; and (c) changes in the energy use of non-participants as a result of utility programs, whether direct (e.g., utility program advertising) or indirect (e.g., stocking practices such as (b) above or changes in consumer buying habits)." Participant spillover is described by (a), and nonparticipant spillover, by (b) and (c). Some parties refer to non-participant spillover as "free-drivers." (TecMarket Works Team, 2006)

Some evaluators subdivide participant spillover into "inside" and "outside" spillover. Inside spillover occurs when, due to the project, additional actions are taken to reduce energy use at the same site, but these actions are not included as program savings. Outside spillover occurs when an actor participating in the program initiates additional actions that reduce energy use at other sites that are not participating in the program.[26]

---

[26] It is worth noting that one implication of all of these definitions is that how a piece of savings is classified may depend in part on the objectives of the program and what outcomes the program has chosen to track. As a key example, program influence achieved through the provision of technical information (henceforth called information-induced savings for shorthand) is clearly a legitimate source of savings, but, depending on the specifics of the situation, could end up being classified either as in-program savings, participant spillover, or non-participant spillover. If the provision of information is considered sufficiently central to the program objectives for the program to directly track this outcome, then information-induced measures may be classified as in-program savings. If information-induced measures are not tracked but are adopted by participants who also adopted rebated measures, and thus entered the tracking system, then they may end up being classified as participant spillover. If untracked information-induced measures are adopted by end-

Because causality is inherent in the very definition of spillover, the spillover savings are inherently net.

The Guidelines further explain that free ridership and spillover are captured in the net-to-gross (NTG) ratio to reflect the degree of program-induced actions. Specifically, the gross energy savings estimate, refined by the realization rate, is adjusted to reflect the negative impacts of free ridership and the positive impacts of spillover. Equation 1 illustrates this adjustment.

$$\text{NTG ratio} = (1\text{-Free ridership}) + \text{Spillover} \qquad \text{Equation 1}$$

Clearly, ignoring spillover results in a downward bias in the NTG ratio.

While underscoring the importance of spillover and supplying important methodological references, the Guidelines provide no additional guidance for estimating spillover. The goal of this Appendix is to provide these general guidelines for estimating both participant and nonparticipant spillover. [27]

2. Key Decisions for Evaluators

Before evaluators decide to estimate spillover, they must make a number of critical decisions:

a. Will the evaluation address participant spillover, nonparticipant spillover, or both?
b. Does the size of the expected savings warrant the expenditure of evaluation funds needed to estimate these savings at an appropriate level of reliability?
c. Which of the two levels of methodological rigor discussed in these guidelines, *standard* or *enhanced*, should be used?
d. Will spillover be estimated based on data collected from end users, those upstream from end users (e.g., vendors, installers, manufacturers, etc.), or both?
e. What is the level of aggregation? Although participant spillover is always estimated at the program level, if an evaluator is attempting to estimate nonparticipant spillover, will the evaluator estimate it at the program level or the market level? One potential reason for estimating nonparticipant spillover at the market level is that, in some circumstances, reliably teasing out the spillover savings attributable to one specific program among many may be nearly impossible due to the difficulty nonparticipants may have in attributing any of their installations to a specific

---

users who did not also adopt a measure for which they received a rebate, then they may be classified as non-participant spillover. While all of this suggests that the precise meaning of these terms can be somewhat specific to the situation, this document is intended to provide methodological guidance that is resilient in the face of such distinctions.

[27] While the spillover guidance provided in this Appendix focuses entirely on estimating benefits, PAs should not forget that they must also estimate the incremental costs associated with each spillover measure. Both the benefits and costs of spillover measures must be included in the total resource cost (TRC) test and the societal test.

program. In such a case, evaluators can choose to conduct market effects studies which include naturally occurring adoptions, program-rebated adoptions, participant and nonparticipant spillover, other program effects that cannot be reliably attributed to a specific program (e.g., upstream lighting programs and the effects of the portfolio of programs on such things as increases in the allocation of shelving space to efficient measures), and other non-program effects due to such factors as DOE Energy Star, programs funded by the American Recovery and Reinvestment Act (ARRA) and the gradual non-program induced evolution of the market in terms of attitudes, knowledge and behavior regarding energy efficiency. The net savings resulting from market effects studies must be included in the portfolio-level benefits-costs analyses.

f. If an evaluator decides to conduct a market effects study, then they must decide whether the study should be focused on the region targeted by a given PA, multiple regions or even the entire state.

Once these questions are answered, evaluators can then use these guidelines in estimating spillover.

3. Program-Specific Methods

3.1. Level of Rigor

Various types of spillover can be estimated using data collected from participating and nonparticipating end users and from participating and/or nonparticipating market actors upstream from the end users (e.g., vendors, retailers, installers, manufacturers). These savings can also be estimated at varying levels of methodological rigor. Program administrators should propose whether a given spillover analysis should receive *standard* or *enhanced* treatment. DPS will review PA proposals and make a determination based on the value of the data balanced against the cost of the research. The primary criterion for whether a given spillover analysis is subject to standard vs. enhanced requirements is the expected magnitude of spillover savings. Factors that the PAs should consider in making its proposal and that DPS staff will consider in making a determination regarding expected magnitude of spillover savings include among others:

1. Past results for the same PA program
2. Program theory or market operations theory
3. National research literature for similar programs.
4. Size of the program
5. Size and complexity of the market
6. Nature of the technology(ies) promoted by the program

Table 1 presents the standard and enhanced levels of rigor for estimating both gross spillover savings and program influence for both end users and those upstream from the end users.

Table 1. Level of Methodological Rigor for Estimating Gross Spillover Savings and Program Influence

	Standard Rigor	**Enhanced Rigor**
Overall Methodological Approach	May rely solely on self-reports from end-users and upstream market actors to support estimates of gross savings or program influence.	Basic self-reports from end-users and upstream market actors typically not sufficient as sole method to support estimates of gross savings or program influence
Estimation of average gross savings for spillover measures for end users (participants and/or nonparticipants).	Simplifying assumptions may be made, such as average gross unit savings being the same for spillover measures as for in-program measures.	Average gross unit savings for spillover measures must be documented empirically, based on a combination of self-reports and/or on-site visits.
Estimation of gross savings from upstream actors (participants and/or nonparticipants).	Self-reports generally sufficient.	Researchers must attempt to confirm self-reports using methods such as changes in sales, stocking or shipment data, review of planned or completed project or permits, or on-sites.
Estimation of program influence for end users (participants and/or nonparticipants).	Basic self-reports generally sufficient.	Enhanced self-reports generally sufficient[28].
Estimation of program influence for upstream actors (participants and/or nonparticipants).	Basic self-reports generally sufficient.	Either additional methods such as quasi-experimental design, econometric analysis, or Delphi panels[29] should be deployed or a case should be made that such methods are either not viable or not cost-effective.
Documentation of causal mechanisms	Recommended but not required.	Required, using methods such as self-reports from end-users or market actors regarding the manner in which the program influenced their behavior, and/or theory-driven evaluation practices. [30]

[28] Basic self-reports typically involve interviewing one participant decision-maker or market actor. Enhanced self-reports on the other hand typically involve more intensive data collection and analysis in the estimation of the net-to-gross ratios. For example, it can include collecting data from more than one participant decision-maker as well as from others such as relevant vendors, retailers, installers, architectural and engineering firms, and manufacturers. It can also include the consideration of past purchases and other qualitative data gleaned from open-ended questions.

[29] Delphi panels can be useful as long as members are provided sufficient market-level empirical data to inform their deliberations. Delphi panels should not be confused with brainstorming.

[30] Documentation of causal mechanisms can include verification of the key cause and effect relationships as illustrated in the program logic model and described in the program theory. Weiss (1997, 1998) suggests that a theory-driven evaluation can substitute for classical experimental study using random assignment. She suggests that if predicted steps between an activity and an outcome can be confirmed in implementation, this matching of the theory to observed outcomes will lend a strong argument for causality: "If the evaluation can show a series of micro-steps that lead from inputs to outcomes, then causal attribution for all practical purposes seems to be within reach" (Weiss 1997, 43).

3.2. Double Counting

PAs should propose methods to avoid double counting both participant spillover and nonparticipant spillover. For example, some participant or nonparticipant spillover measures might have received assistance (information and/or incentives) from some other PA programs. In such cases, measures receiving assistance from other PA programs should be subtracted for the spillover estimates. Or, in other cases, two programs could be targeting the same market for the same measures. In such cases, because it would be challenging to accurately allocate spillover savings attributable to each program, expert judgment may be used. Under no circumstances, when the possibility of double counting exists, should a PA claim the sum of the spillover savings separately estimated for each program without making the appropriate adjustments. Determining how the estimated spillover savings should be allocated among different programs within a given PA's portfolio and/or across PA portfolios can be based on such factors as the size of the program budgets, program theories and logic models that demonstrate the causal mechanisms that are expected to produce spillover, and the results of theory-driven evaluations (Weiss, 1997; Donaldson, 2007).

3.3. Calculation of the Program-Level Spillover Rate

While PAs are free to calculate spillover rates in a variety of ways, there must be at least one method that is used consistently across all PAs. The formulation of a NTG ratio presented in the Guidelines is repeated in Equation 2:

$$\text{NTG Ratio} = (1 - \text{Free Ridership}) + \text{Spillover} \qquad \text{Equation 2}$$

Equation 2 illustrates that the spillover rate is added to 1-Free Ridership to produce the NTGR. Given the additive nature of the spillover rate in Equation 2, the spillover rate must be calculated as in Equation 3:

$$\text{Spillover Rate} = \frac{\text{Net PSO+ Net NPSO}}{\textit{Ex Post}\ \text{Gross Program Impacts}} \qquad \text{Equation 3}$$

4. Estimating Spillover at the Market Level

In some cases, it might not be possible to reliably estimate nonparticipant spillover at the program level due to multiple program interventions in the same market involving multiple market actors. In such cases, market effects studies can be done for specific measures and markets, e.g., commercial HVAC. This Appendix does not provide any guidelines for conducting such studies but rather refers evaluators to other sources such as Eto, Prahl, and Schlegel (1996), Sebold et al. (2001) and TecMarket (2005).

5. Sampling and Uncertainty

Sampling for both program-level and market level spillover studies should be done in accordance with the *Sampling and Uncertainty Guidelines* in Appendix E.

6. Levels of Confidence and Precision

As discussed in the main body of the DPS guidelines, the minimum standard for confidence and precision for overall net savings at the program level is 90/10. Here, overall net savings includes both in-program net savings and any reported spillover savings. The achieved level of confidence and precision for overall program net savings must be reported at the 90% level of confidence.

If reported savings results include spillover savings, there is no required level of confidence and precision specifically for the individual components of net savings from in-program measures and net savings from spillover. However, PAs are still accountable for achieving 90/10 for overall program net savings. The standard error of overall program-level net savings can be calculated by combining the achieved levels of confidence and precision for the net savings from in-program measures and for spillover savings using standard propagation of error formulas (Taylor, 2006; TecMarket, 2004).[31] While there are no precision *requirements* for the individual components of net savings from in-program measures and the net savings from spillover measures, the precision actually achieved for each of these components should be reported at the 90% level of confidence, in order to help facilitate assessment of the reliability of the results.

7. Deemed Approaches

Of course, there might be situations in which all key stakeholders are willing to agree that spillover is not zero but the expense to estimate it reliably is prohibitive. In such cases, a PA may negotiate a deemed spillover rate based on a review of the literature and the program theory and logic model that together describe reasonably well the causal mechanism that is expected to generate spillover.

---

[31] This is generally true as long as each of the individual components making up the total net savings estimate (e.g., gross savings, free riding, spillover, etc.) has been estimated based on independent random samples and methods that allow for the calculation of standard errors. However, there are legitimate circumstances under which the sample designs and methods for one or more components do not meet these requirements. One example is a market effects study in which total net program impacts are estimated using a preponderance of evidence approach. Another example (some aspects of which are discussed in the next section) is a case in which one or more components are deemed. A third example is a case in which multiple methods are used to estimate net impacts or the net-to-gross ratio, and a Delphi analysis is used to integrate the results. If *none* of the individual components meet these requirements, then clearly the issue of precision does not apply. If some components meet these requirements but others do not, then the program administrator should take clear note of this fact and propose an approach to ensuring that the components of the study that do meet these requirements are performed in a manner that gives due attention to limiting the effects of sampling error.

8. References

California Public Utilities Commission: Energy Division and the Master Evaluation Contractor Team. (2007). *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches.*

Donaldson, Stewart I. (2007). *Program Theory-Driven Evaluation Science: Strategies and Applications*. New York: Psychology Press.

Eto, Joseph, Ralph Prahl and Jeff Schlegel. (1996). *A Scoping Study on Energy-Efficiency Market Transformation by California Utility DSM Programs*. Prepared for The California Demand-Side Measurement Advisory Committee: Project 2091T

Frederick D. Sebold, Alan Fields, Shel Feldman, Miriam Goldberg, Ken Keating and Jane Peters. (2001). *A Framework for Planning and Assessing Publicly Funded Energy Efficiency: Study ID PG&E-SW040.* Prepared for the Pacific Gas & Electric Company.

Taylor, John R. (1997). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Sausalito, California: University Science Books.

TecMarket Works Team. (2004). *The California Evaluation Framework.* Prepared for the California Public Utilities Commission and the Project Advisory Group, Framework

The TecMarket Works Team. (2006). *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals.* Directed by the CPUC's Energy Division, and with guidance from Joint Staff.

Weiss, Carol H. (1997).Theory-Based Evaluation: Past, Present, and Future. In Debra J. Rog and Deborah Fournier (eds.) *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods*. San Francisco: Jossey-Bass Publishers.

Weiss, Carol H. (1998). *Evaluation*. Upper Saddle River, New Jersey: Prentice Hall.