H Gil Peach & Associates/Scan America®

Verification Review of Program Year 2014 Evaluation Results

Report for the Nova Scotia Utilities and Review Board

H. Gil Peach, Ph.D.
John Mitchell, B.S.
March 26, 2015

Scan America® Vision Statement

To be a world leader in developing truthful measurement and useful results; to support development of efficient, ethical, and effective practices, sustained economically; to advance human development.

Scan America® Goals Statement

- Excellence in the integration of knowledge, method, and practice
- Improvement and learning at all levels
- Contextually sound measurement, analysis, and reporting
- Anticipate and meet the needs of our clients
- Awareness of human relevance and of the ethical core of research
- To go further, to find better ways

Scan America® Mission Statement

With extensive experience in North America, Europe, and Asia, together we can provide the full range of management, planning, and evaluation services – wherever and whenever there is a need.

H Gil Peach and Associates
Website
www.scanamerica.net

Scan America®

H. Gil Peach & Associates LLC

16232 NW Oak Hills Drive
Beaverton, Oregon 97006-5242, USA

Telephone: (503) 645-0716 EIN: 93-1323715
Fax: (503) 946-3064

H. Gil Peach, Ph.D.
hgilpeach@scanamerica.net

Scanada Consultants Limited
1446 Birchwood Drive
Mississauga, Ontario L5J 1T2, Canada

C. Eric Bonnyman, M.Sc.
scanada@sympatico.ca

WSP Sverige AB
SE-121 88 Stockholm-Globen, Sweden
Visiting address: Arenavägen 7
Phone: +46-(0)10 722 50 00

Agneta Persson, M.Sc.
agneta.persson@wapgroup.se

Verification Review

Of Program Year 2014 Evaluation Results

Table of Contents

Figures

Tables

I. Overview

This report is a savings verification review conducted by H. Gil Peach & Associates, LLC for the Nova Scotia Utility and Review Board (Board). It reports on verification of electricity energy savings and demand reduction estimated by Econoler, the Independent Evaluator for the Demand Side Management Administrator's (DSM Administrator's) calendar 2014 energy savings and demand reduction programs.

The primary purpose of this report is to review, verify, and if appropriate, recommend adjustments to the savings data and estimates created by evaluation consultants engaged by the DSM Administrator.[1] For this reason it is focused on impact evaluation. The report also provides limited comments on other parts of the evaluation.[2]

The examination is based on careful review of the evaluation for each program component or initiative in the DSM portfolio, including review of each evaluation's explicit or implicit design, of evaluation methods and standards, and/or evaluation protocols used in each evaluation. It is also based on review of the program data tracking system, on discussions with staff and delivery agents concerning program procedures and methods for program Quality Assurance/Quality Control (QA/QC) and selected "due diligence" site visits to check installation counts and quality of work. In addition, the savings verification consultant watches the evolving general standards of practice in DSM and DSM evaluation to highlight "better practices". This is a verification review of the set of evaluations of Efficiency Nova Scotia's programs and pilots for Program Year 2014.

The primary purpose of this report is to review, verify, and if appropriate, recommend adjustments to the savings data and estimates created by the independent evaluation consultants engaged by the DSM administrator

[1] Nova Scotia Power, Inc. served as the DSM Administrator for calendar years 2008 and 2009. Responsibility transferred to Efficiency Nova Scotia Corporation in October 2010. Calendar year 2011 represented the first full program cycle with Efficiency Nova Scotia as the DSM Administrator and 2014 is the seventh program year from the beginning of the effort in 2008. As a rule of thumb, after the fifth year a major effort such as this is considered mature.

[2] For example, process and market evaluation.

II. Demand-Side Management

It can be useful to understand DSM measurement and evaluation in context by reviewing how DSM works. Probably the most important thing is to see DSM measurement and evaluation as a single activity, but as a *cycle of repeated activities subject to continuous process improvement*. Figure 1 emphasizes the DSM cycle with four main parts designated by the larger circles. There are also sub-parts for each of the primary parts. The sub-parts for Planning are shown off to the left side of the figure). Each main part contributes knowledge to improve the other main parts from cycle to cycle.

At a formal level, the evaluator provides measurements that can feed back knowledge to the programs both as they are running and for the next program cycle. But, it is also true that direct experience in conducting programs creates knowledge of products, markets and engagement that the evaluators do not have, so new knowledge and insight s flow both ways in all parts of the diagram. Experienced program managers acquire deep practical knowledge of specific sets of markets, technologies, codes and practices and engagement. For example, when management and staff share information with the independent evaluator, the evaluations improve.

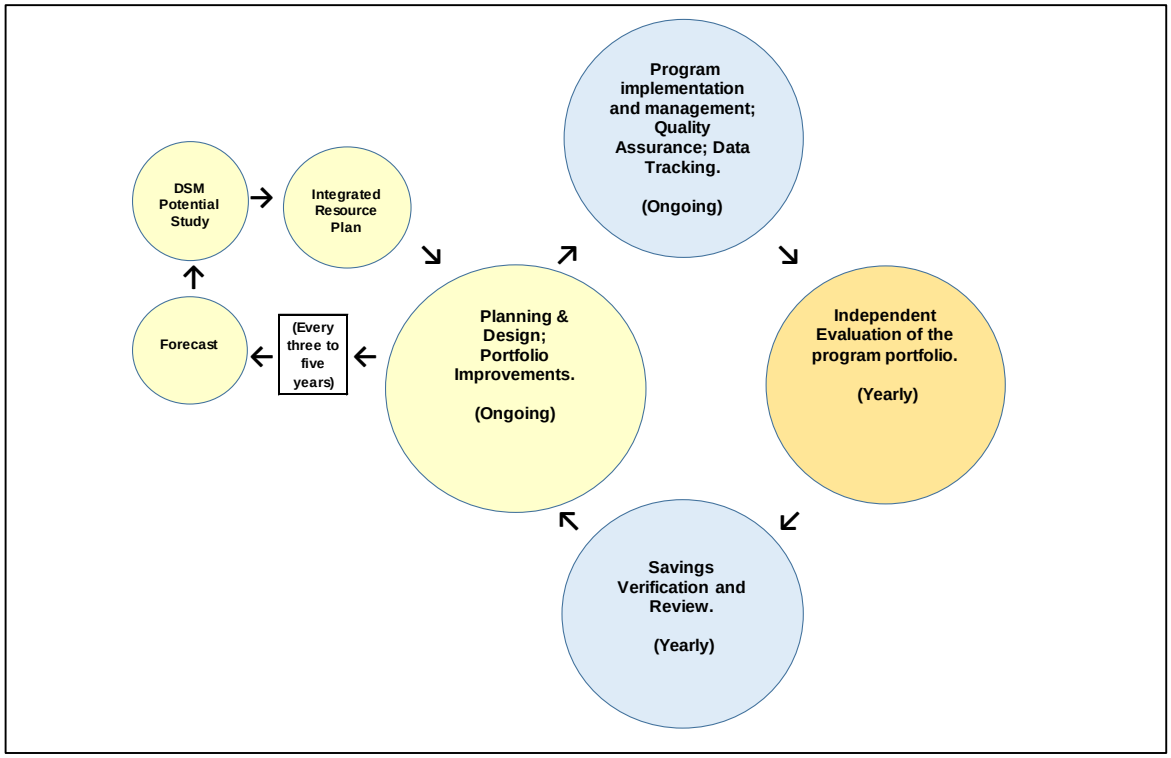


Figure 1: The Demand-Side Management Cycle.

The arrows in Figure 1 show the direction of the primary flow of activity within a DSM

cycle. Each step in the Demand-Side Management Cycle feeds information forward to the Planning step, and planning proceeds continuously. In addition, every three to five years there is an intensified major planning effort that sets the pattern for developing the program portfolio until the next major planning effort. (This is shown off to the left in Figure 1, and consists of the development of forecasts, a potential study with technical, economic and achievable potential and an integrated resource plan).[3]

The Implementation step in each cycle feeds information forward to Planning through the records in the program data tracking system and the development of practical knowledge gained by program managers and delivery agents. The Plan provides guidance, but in Implementation one is "on the point" encountering emergent realities that require alert attention and quick adaptation based on new information. In this way, the programs are adapted from planning perspectives to field realities in order to move programs forward through (usually incremental) changes to adapt to situations, forces and limits that the planning functions could not have anticipated.[4]

The Evaluation step should also feed timely information back to the Program Implementation step in order to provide opportunity for in-progress program corrections within the same cycle (without waiting one, two or three years for full evaluation reports). However, Evaluation is primarily directed toward providing independent program assessments for the next Implementation step and for portfolio planning.

Both the Program Implementation step and the Evaluation step feed information forward to the Savings Verification and Planning steps, for example, through the results of electrical measurements. The Evaluation step adds results of independent measurements of auxiliary variables from direct measurement and results from observations, site inspections, surveys, interviews, literature reviews and analysis/calculations (for example, of electrical measurements, hours of operation, interaction effects, and the assessment of free-riders and spillover effects).

---

[3] Morse, William L. & H. Gil Peach, "Control Concepts in Conservation Supply," Energy, Vol. 14, No. 11, Pp. 727-735, 1989; Gellings, Clark W. & John H. Chamberlin, *Demand-Side Management Planning Concepts & Methods*, Second Edition, Liburn Georgia: The Fairmont Press, 1992; Gellings, Clark W. & John H. Chamberlin, *Demand-Side Management Planning*, Liburn, Georgia: The Fairmont Press, 1993.

[4] For an early article on this point, see Keating, Kenneth M., Ruth L. Love, Terry V. Oliver, H. Gil Peach & Cynthia B. Flynn, "The Hood River Project; Take a Walk on the Applied Side," Pp. 112-118, *The Rural Sociologist*, Vol.5, No. 2, May 1985.

III. The Program Portfolio and First Year Results

There are seventeen program components or initiatives for calendar 2014. Table 1[5], first column, shows first year net energy savings at generator as reported by the evaluator, Econoler, for program components and initiatives for 2014. Overall, the Econoler reported a total of 151,888 GWh first year net energy savings. In addition, Codes and Standards is reported to have achieved net electricity savings of 31.433 GWh at the generator. The overall first year net peak demand reduction (Table 2) is reported at 27.077 MW at the generator. In addition Codes and Standards is reported to have achieved a net first year peak demand reduction of 5,164 MW at the generator

A. First Year Results by Component or Initiative

First year energy savings are reported in Table 1. First year demand reductions are shown in Table 2. [6]

- **First Year Energy Savings**. As can be seen in Table 1, component performance is spilt about 50/50. First year energy savings targets were exceeded by seven (7) program components or initiatives: Home Energy Assessment, Green Heat, Low Income Homeowner, Residential Direct Install, Home Energy Report, Custom Retrofit, and Custom New Construction. Seven (7) programs or initiatives underperformed: Appliance Retirement, Instant Savings, New Home Construction, Tenant Units – Rental Properties and Condos Service, Solar, Business Energy Rebates, and Business Energy Solutions. Three (3) components or initiatives did not have targets: Residential Financing, LED Holiday Lighting Exchange and Codes & Standards.

- **First Year Net Demand Reductions.** First year net demand reduction results (Table 2) showed a similar pattern: Seven (7) of the program components or initiatives exceeded their targets and seven (7) underperformed. Three did not have targets: Residential Financing, LED Holiday Lighting Exchange and Codes & Standards.

.

[5] Table 1 is based on Table 6, Page 12 of the Econoler Evaluation of 2014 DSM Programs Executive Summary.

[6] Table 2 is based on Table7, Page 13 of the Econoler Evaluation of 2014 DSM Programs Executive Summary.

Table 1: First Year Net Energy Savings at the Generator.

DSM COMPONENT	FIRST YEAR				
	EVALUATED	PLANNING TARGETS		TRACKING SYSTEM	
	Net Savings (GWh)	Target GWh Net Energy Savings	Net as % of Target	Tracked GWh Net Energy Savings	Net as % of Tracked
Residential					
Appliance Retirement	3.941	4.1	96.1%	3.941	100.0%
Instant Savings	9.360	13.4	69.9%	15.026	62.3%
Home Energy Assessment	5.785	4.8	120.5%	5.785	100.0%
Green Heat	1.472	1.3	113.2%	1.580	93.2%
Low Income Homeowner	2.311	1.9	121.6%	2.311	100.0%
New Home Construction	4.572	5.9	77.5%	4.572	100.0%
Residential Direct Install	24.607	23.0	107.0%	24.960	98.6%
Rental Properties and Condos Service-Tenant Units	2.720	3.0	90.7%	2.720	100.0%
Solar	0.062	0.2	31.0%	0.062	100.0%
LED Holiday Light Exchange	0.192	No Target	No Target	0.246	78.0%
Home Energy Report	25.437	15.4	165.2%	Not Tracked	Not Tracked
Residential Subtotal	80.459	72.9	110.4%	61.203	131.5%
BNI					
Business Energy Rebates	24.610	27.0	91.1%	23.911	102.9%
Custom Retrofit, EBC, EMIS	26.988	19.3	139.8%	27.314	98.8%
Custom New Construction	6.656	4.6	144.7%	7.122	93.5%
Business Energy Solutions (Note 2)	13.175	14.0	94.1%	12.372	106.5%
BNI Subtotal	71.429	64.9	110.1%	70.719	101.0%
Portfolio Total	151.888	137.8	110.2%	131.922	115.1%
Codes & Standards					
Codes & Standards	31.433	No Target	No Target	Not Tracked	Not Tracked

Note 1: The 2014 evaluation for Appliance Retirement, Home Energy Assessment, Low Income Homeowners, New Home Construction, Tenant Units for Rental Properties and Condos Service and Solar did not involve a full-scale impact evaluation, but rather a validation of ENSC's 2014 tracked savings based on the parameters obtained from the 2013 evaluation and a reassessment of some key parameters.
Note 2: Rental Properties and Condos Services common spaces measures are claimed under Business Energy Solutions; Residential Financing does not have separate savings information.

Table 2: First Year Net Demand Reduction at Generator.

DSM COMPONENT	FIRST YEAR				
	EVALUATED	PLANNING TARGETS		TRACKING SYSTEM	
	Net Savings (MW)	Target MW Net Reduction	Net as % of Target	Tracked MW Net Reduction	Net as % of Tracked
Residential					
Appliance Retirement	0.627	0.7	89.6%	0.627	100.00%
Instant Savings	1.479	2.1	70.4%	2.149	68.82%
Home Energy Assessment	1.423	1.2	118.6%	1.423	100.00%
Green Heat	0.594	0.4	148.5%	0.705	84.26%
Low Income Homeowner	0.539	0.4	134.8%	0.539	100.00%
New Home Construction	1.175	1.5	78.3%	1.175	100.00%
Residential Direct Install	5.734	5.4	106.2%	5.814	98.62%
Rental Properties and Condos Service - Tenant Units	0.634	0.7	90.6%	0.634	100.00%
Solar	0.000	0.1	0.0%	Not Tracked	Not Tracked
LED Holiday Light Exchange	0.414	No Target	No Target	1.228	33.71%
Home Energy Report	2.892	3.6	80.3%	Not Tracked	Not Tracked
Residential Subtotal	15.511	16.0	96.9%	14.294	108.51%
BNI					
Business Energy Rebates	4.880	4.8	101.7%	5.728	85.20%
Custom Retrofit, EBC, EMIS	1.599	3.6	44.4%	1.885	84.83%
Custom New Construction	2.517	1.1	228.8%	2.619	96.11%
Business Energy Solutions (Note 2)	2.570	2.5	102.8%	3.218	79.86%
BNI Subtotal	11.566	11.9	97.2%	13.450	85.99%
Portfolio Total	27.077	27.9	97.1%	27.744	97.60%
Codes & Standards					
Codes & Standards	5.164	No Target	No Target	Not Tracked	Not Tracked

Note 1: The 2014 evaluation for Appliance Retirement, Home Energy Assessment, Low Income Homeowners, New Home Construction, Tenant Units for Rental Properties and Condos Service and Solar did not involve a full-scale impact evaluation, but rather a validation of ENSC's 2014 tracked savings based on the parameters obtained from the 2013 evaluation and a reassessment of some key parameters.
Note 2: Rental Properties and Condos Services common spaces measures are claimed under Business Energy Solutions; Residential Financing does not have separate savings information.

As shown in Figure 2, a small number of programs produce most of the energy savings. In this figure, the programs from Table 1 are listed on the left from top to bottom in order of ranking by amount of energy savings produced. Codes & Standards, Custom Retrofit, Business Energy Rebates, Home Energy Report and Residential Direct Install are the top producers of first year conserved GWh. Code & Standards are considered to be a DSM initiative outside the portfolio. For the portfolio, the top four programs produce 103.65 GWh or about 68% of the 151.888 GWh net portfolio energy savings.

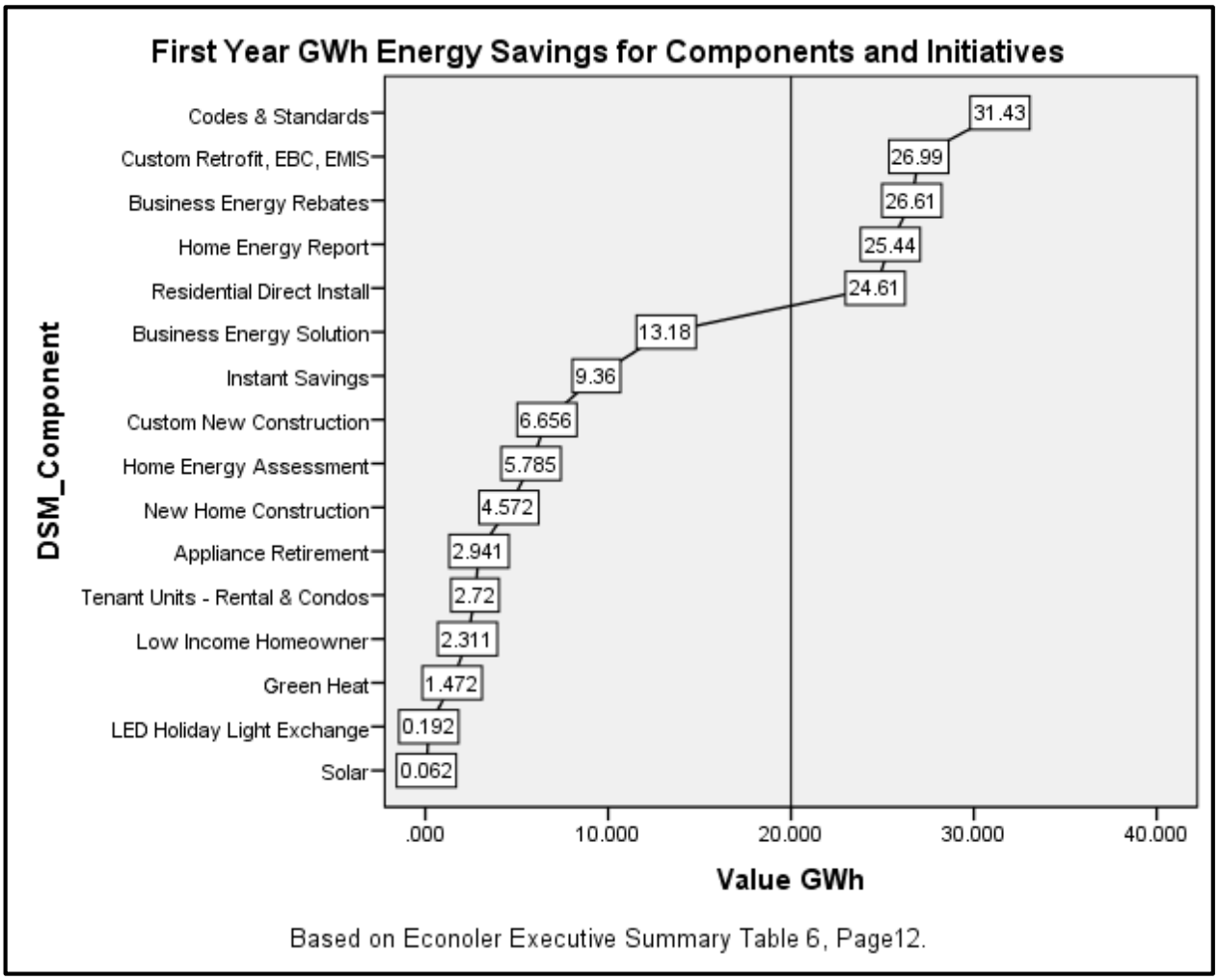


Figure 2: Rank Ordered First Year GWh Contribution.

In Figure 3, the programs are sorted by contribution by amount of first year demand-reduction. Values of MW reduction are shown on the graph and the values are plotted in rank order. The order of programs on contribution to demand reduction is not the same as the order of the programs by energy savings, but the pattern is much

the same. The Residential Direct Install program, Codes & Standards, Business Energy Rebates, Home Energy Report, Business Energy Solutions and Custom New Construction top producers. Codes & Standards is not considered part of the portfolio. The top five portfolio programs produce 18.233 MW or about 67% of the 27.077 MW first year portfolio demand reduction.

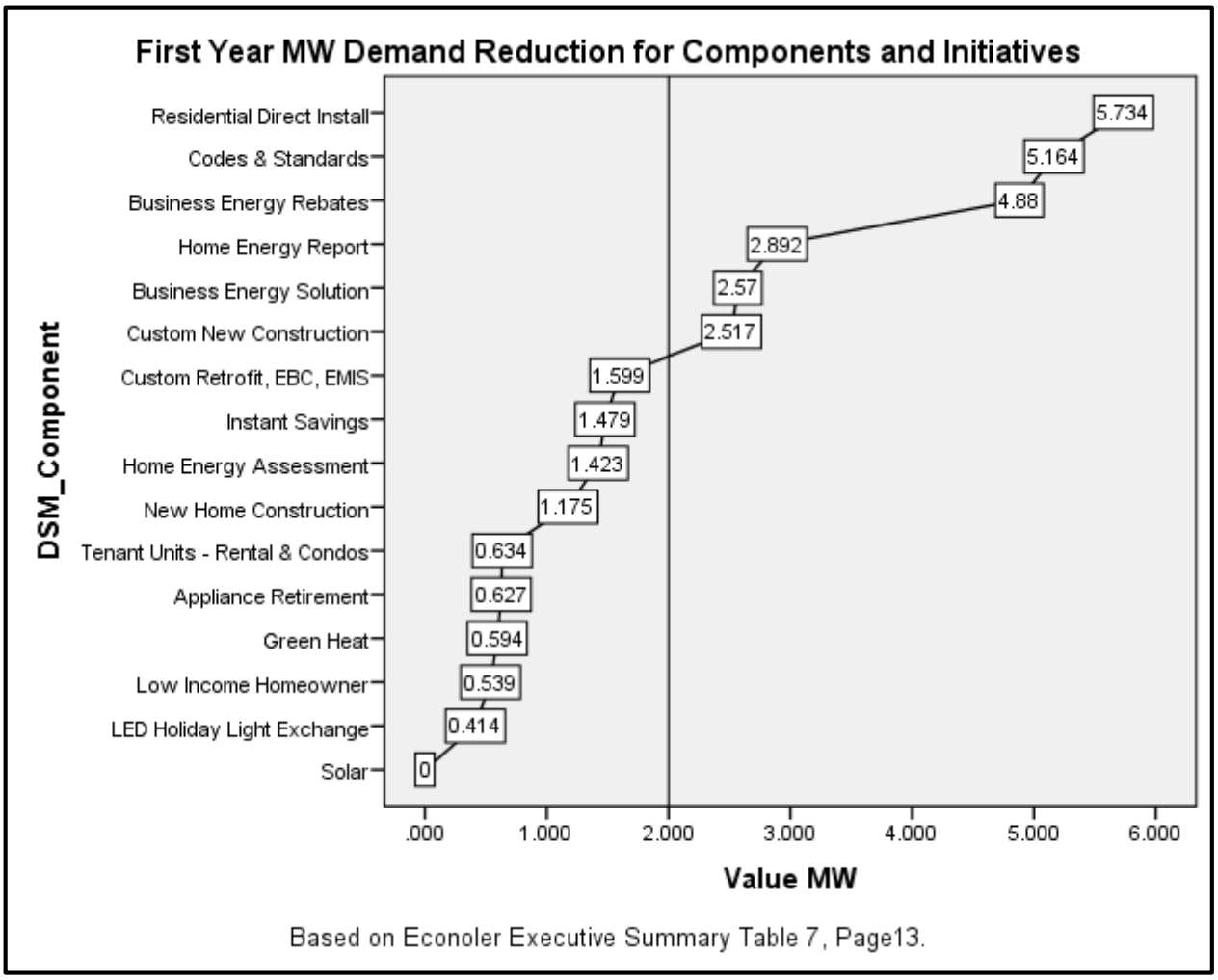


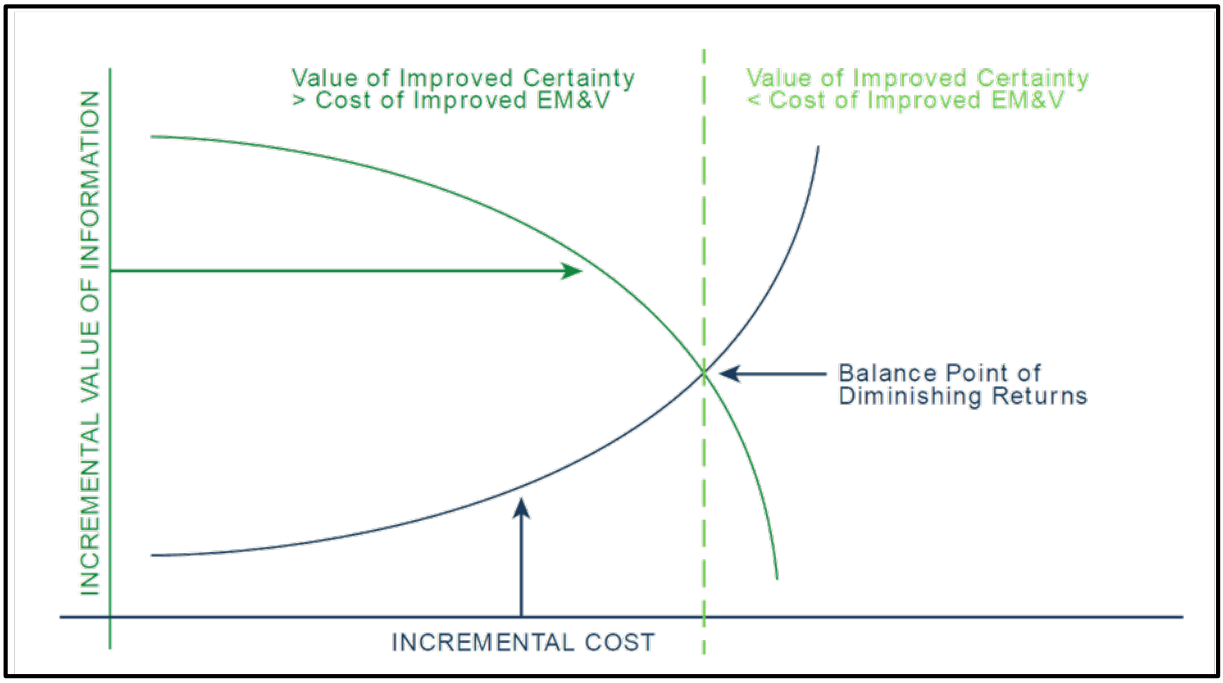
Figure 3: Rank Ordered First Year MW Demand Reduction Contribution.

B. A Better Analysis

Currently, evaluations are required to produce the level of informational results shown above in Tables 1 & 2 and in Figures 2&3. Impact evaluations are to produce first year energy savings and first year demand reductions as primary results. The evaluator is not required to do more, but could be required to take the next step in analysis.

The value of directing evaluators to develop a more complete picture for future evaluations would have to be considered in terms of the incremental cost. A conventional economics conceptual picture of this relation is provided in Figure 4.[7] I

Figure 4: Evaluation Benefit and Cost.



There are two important reasons why developing a better indicator than first year result is worth the effort:

(1) **Time Dimension.** First, there is a time dimension to the production of energy savings and demand reductions that is missed entirely when first year savings is used as the sole indicator of energy savings results of DSM programs. Currently, the evaluations produce savings as a flat first year figure, as if every

[7] Figure 4 is a reproduction of Figure 7.1 Incremental Value of Information vs. Incremental Cost of Evaluation, P. 7-13 in State and Local Energy Efficiency Action Network, *Energy Efficiency Program Impact Evaluation Guide,* 2012. Prepared by Steven R. Schiller, Schiller Consulting, Inc., www.seeaction.energy.gov.

conserved kWh had the same value. The value of incremental information in moving the indicators to also include life cycle savings and a graphical illustration of how conserved kWh occur over years is that this will immediately differentiate programs with a strong multi-year performance forward in time from programs that simply produce kWh savings today that will quickly disappear. This differentiation can be illustrated by looking at what appear to be the best performing programs in size of contribution to first year energy savings. For the 2014 program evaluation these are:

- Codes & Standards,
- Custom Retrofit,
- Business Energy Rebates,
- Home Energy Report and
- Residential Direct Install.

But, of the five components or initiatives that look to be performing similarly as top savings producing programs in the first year, four have substantial energy savings year after year into the future and one does not. Because the source of energy savings for the four programs is technical improvement or a combination of technical improvement plus regulation and the fifth is a standalone behavioral program (Home Reports), only the four technical and technical/regulatory programs can be expected to continue to produce reliable savings for many years.

Developing a savings weighted measure life indicator system to serve as a primary indicator of energy savings would immediately differentiate programs according to the time dimension of their energy savings. *The first year savings indicator, by contrast, is a flat analysis. Its defect is that is cannot differentiate ephemeral behavioral standalone programs from technical and technical/regulatory programs.* With only the first year saving as the indicator of impact, short term behavioral standalone program kWh appear to have equal value to technical and technical/regulatory kWh as an artifact of the indicator system. Addition of the time dimension crisply separates standalone behavioral programs from technical and technical/regulatory programs.

This separation can be further developed using a program classification for DSM that was originally developed for the broader environmental area. According to Heberlein [8], programs can be usefully classified into three types:

[8] Heberlein, Thomas A., Navigating Environmental Attitudes. Oxford & New York: Oxford University Press, 2012. See, in particular, Table 1.1, Attitude and the Three Fixes for Environmental Problems, Page 9.

(a) Purely or Primarily Technical. Most traditional DSM programs are implemented directly as purely technical or engineering projects. Of course, all programs have behavioral aspects. These are usually handled not as if they were separate standalone DSM programs but as ancillary informational, marketing and promotional (including motivational) efforts to implement technical programs. An environmental example (given by Heberlein) is the construction of water projects for flood control, which, while they have social impact and social components, are *primarily engineering projects*.

A Demand Side Management example would be an economic retrofit in a manufacturing setting or installation of insulation in a house. According to Heberlein, *purely technical projects require consistency with dominant attitudes and values in the social context in which they are introduced.*

(b) Purely or Primarily Cognitive. Other programs can be informational or promotional. The naive version of this program type is purely informational. More advanced versions can have behavioral aspects.

- An environmental example given by Heberlein is flood plain maps, which, if the "rational man" hypothesis at the heart of conventional economics were true, would automatically result in no building on flood plains.

- A Demand Side Management example would be the provision of informational brochures on energy efficient LEDs.

Of course, in reality, "information only" programs normally have very limited effect in the absence of planning for creative marketing, promotional and motivational campaigns.

(c) Structural/Regulatory. A third approach is structural.

- Heberlein's environmental example is a regulation that there will be no social insurance for banks that finance projects in flood plains. This, in contrast to the information approach, is effective in preventing building in flood plains. Because bank will not assume the lending risk, t*he product cannot be found in the market*.

- A Demand Side Management example would be a code that requires a product to meet an energy efficiency standard so that a

consumer *cannot find the old product in the market*. This type of program strategy in its DSM application was originally developed by Dennis Nelson and others at BC Hydro: building practices are improved though a sequence of incentive and training programs and then the new level of practice is required by law as a minimum standard.[9] Then the process is iterated.

What we now term "market transformation programs" can work this way either by driving inferior products from the market or by accompanying use of market forces with a system of escalating regulatory standards that lock in advances in practice (and then iterate into the future). *According to Heberlein, structural programs must be consistent with dominant public attitudes and values.*

In summary, we assert a classification of three program types: primarily technical, informational and technical/regulatory. They have different levels of effectiveness in practice and are appropriate for different situations. Two of these types produce substantial energy savings forward in time over multiple years and one does not. The types are not equivalent in the time dimension.

(2) Second, *the origins of Demand Side Management are in service to the Resource Planning function* (Figure 5).[10] Originally, "energy conservation programs" were not considered to be Demand Side Management. Just under forty years ago, utilities drew a line at the meter and while always interested in efficiency on the utility side they did not cross over to improve efficiency on the customer side through DSM programs. (They would cross over to help make an important industrial customer more cost effective or to steer large new construction projects and new housing development projects towards electrical energy use. But these were load retention programs and load building programs).

To secure agreement to the concept that "a conserved kWh is equivalent to a generated kWh" required a demonstration that savings could be reliably

[9] At approximately the same time, a different approach to market transformation of technologies was being developed in Sweden, drawing from military procurement strategies (Westling, Hans). Later the more generic form of market transformation and the phrase "market transformation" came into general use for a set of DSM strategies in contrast to "resource acquisition".

[10] Figure 5 is a reproduction of Figure ES-1 Evaluation Objectives, P. xiv in State and Local Energy Efficiency Action Network, *Energy Efficiency Program Impact Evaluation Guide,* 2012. Prepared by Steven R. Schiller, Schiller Consulting, Inc., www.seeaction.energy.gov.

measured and reliably produced. Synonyms for "reliable" are *trustworthy, unfailing,* and *dependable*. Originally, that meant that a DSM program had to be purely or primarily technical.

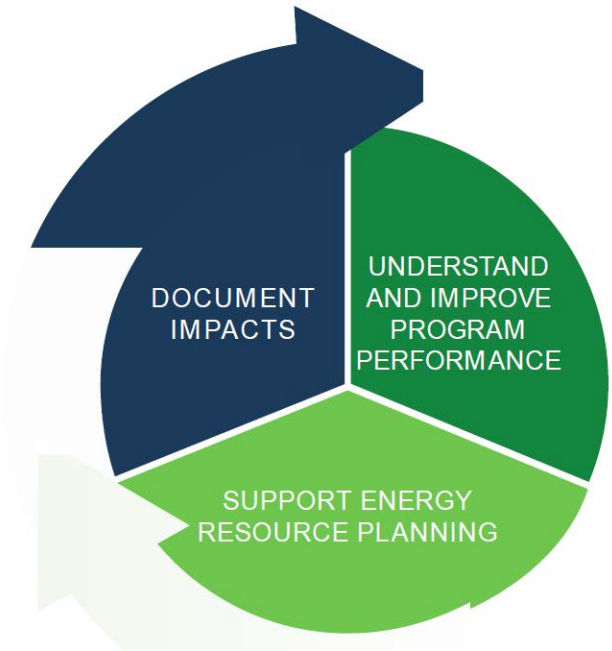


Figure 5: Evaluation of DSM Programs Serves Resource Planning.

In the original resource planning context, when the DSM concept was accepted, a standalone behavioral program would not have been seen as credible. (Although several utilities had successful experience with energy and demand curtailment programs as part of rationing during World War II, these were understood to be emergency programs necessary during wartime and this experience would not have be perceived to be applicable to normal DSM planning.) Adding the time dimension as a primary indicator system by reporting program components or initiatives by savings weighted measure life highlights programs that are reliable and can be counted on over time. Reliable is used in the sense of:

- Trustworthy,
- Unfailing, and
- Dependable.

It differentiates technical and technical/structural programs that provide long-term reliable savings useful in resource planning from standalone behavioral

programs that may provide immediate but ephemeral energy savings.

Typically standalone behavioral programs can be a useful tool during energy crises. For example, in California when markets were opened to competition and regulation was relaxed, a number of credible competitors were attracted but, open markets with weak regulation also invite predatory market entrants and among them, a criminal market player (Enron) was able to game the whole electric system to create artificial supply shortages. The extent of Enron's gaming of the system caused customers to be triaged in order to maintain essential services. This crisis was endured for a few years since electricity markets do not correct themselves. During that time of sustained emergency, emergency behavioral programs coupled with emergency rate designs were often effective in curtailing load. This worked, in part, due to the unusual socioeconomic context – since nearly everyone was attuned to and repeatedly affected by interruptions of service in the continuing crisis, nearly all customers responded to emergency programs.

But emergencies are exceptional times, and resource planning is inherently concerned with at least moderate time durations. To be viable for long-term resource planning, a DSM program's time dimension must be both substantial and unambiguous. System planners, regulators and the parties need to easily see the durations of resource alternatives. *Standalone behavioral programs that yield situational and temporary savings of narrow durations produce essentially "squishy" conserved kilowatt hours, where by "squishy" we indicate:*

- Not firm,
- not steady,
- Not fixed
- Amorphous as to physical causation and not well defined, and
- Not reliable over a reasonable time duration and so not useful.

Standalone behavioral programs may be useful in emergency situations. But it is the primarily technical and technical/regulatory programs that are useful in long-term resource planning.

For these two reasons (the time dimension helps differentiate the relative value of programs and DSM is in the service of resource planning) addition of a systematic indicator for life cycle energy savings is of high value is helping prevent resource acquisition errors – otherwise standalone behavioral programs appear equivalent to primarily technical and technical/regulatory programs as an artifact of using first year

energy savings as the sole impact indicator. From a level of evaluation effort perspective and an evaluation cost perspective, the cost of including the time dimension is negligible, since the calculation is generally performed *ex-ante* in the context of developing a potential study. The model can be easily re-run *ex-post* by the evaluator using actual measures installed and measure counts per program.

As a concrete illustration, Table 3 shows approximate first year savings (GWh) for four program components or initiatives in column one. The second column is the value used for years in the calculation of savings weighted measure life. The final column gives the full GWh production of the (modeled) program.[11] We have here modified the Home Energy Report savings weighted measure analysis (although the terminology is not quite right for this program type since it has an amorphous content so there can be no empirical calculation by individual measure). So, we can assume 50% savings in Year 2, 25% in Year 3, 10% in Year 4 and 5% in Year 5, followed by zero or an average measure life of 1.9 years.[12] These values are arbitrary; reasonable variations would not change the result.

Table 3: First Year vs. Full Energy Savings.

Energy Savings Values			
Component	**First Year GWh**	**SWML2015 (Years)**	**Full GWh**
Home Energy Report	25.437	1.90	48
Business Energy Rebates	26.61	11.74	312
Codes & Standards	31.433	25.00	786
Residential Direct Install	24.607	22.50	554

The values for First Year GWh from Table 3 are used to develop Figure 6. In this figure, the apparent relative value of the Home Energy Reports program is high

[11] These values are estimates and are based on the 2014 evaluation data plus data from the DSM Administrator's savings weighted measure life calculations. Values used here are approximations for purposes of illustration. The DSM Administrator's measure lives are conservative compared with those used by the Northwest Power Planning and Conservation Council.

[12] This is a hypothetical series. Only a few projects of this type have been tracked for annual decay rates which, on the basis of a handful of examples, Khawaja and Stewart report range from 11% to 83%. For a conceptual presentation of developing a decay curve for Home Reports programs, see Khawaja, M. Sami and James Stewart, "Long-Run Savings and Cost-Effectiveness of Home Energy Report Programs," Cadmus, 2014. (http://www.cadmusgroup.com/papers-reports/long-run-savings-cost-effectiveness-home-energy-report-programs/).

because the metric is only first year savings. The Full GWh from Table 3 are used in Figure 7, showing that the apparent relative value of the Home Energy Reports program is very small when the time dimension of energy savings is included in the impact indicator.[13].

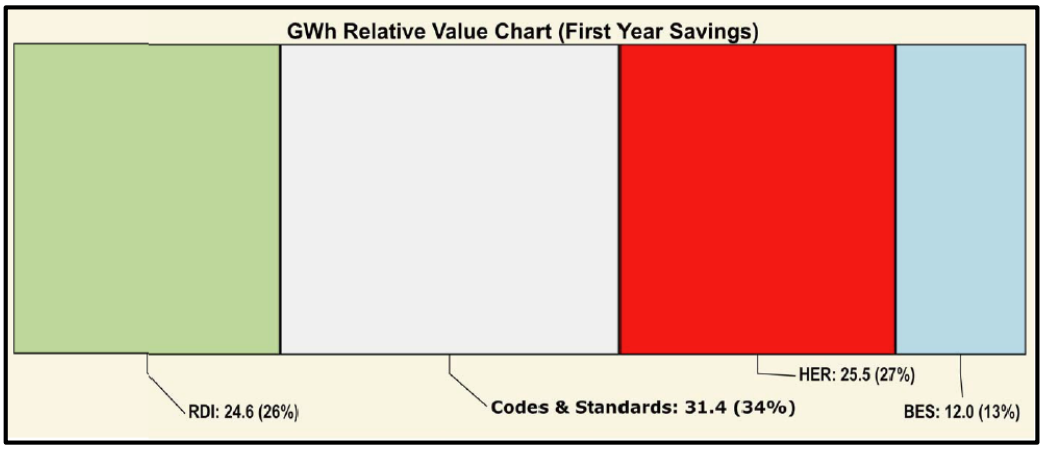


Figure 6: Relative Value of Programs by First Year Energy Savings (approximate).

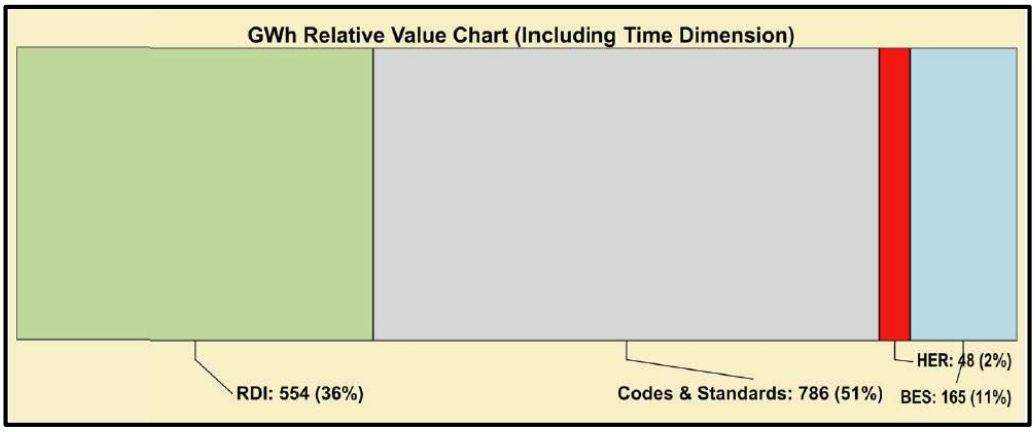


Figure 7: Relative Value of Programs by Full Energy Savings (approximate).

[13] An arbitrary limit of twenty-five years is used in Figure 7. Technically, the measure life for Code and Standards is infinite because the DSM Administrator takes no credit for this program. If the DSM Administrator were assigned credit, the credit would gradually erode on the assumption that the code would eventually be improved without support from the energy efficiency utility. See: Mahone, Douglas, Nick Hall, Lori Megdal, Ken Keating & Rick Ridge, "Codes and Standards White Paper on Methods for Estimating Savings," Prepared for Marian Brown, SCE in support of Statewide NRNC MA&E, April 7, 2015. (http://www.cpuc.ca.gov/NR/rdonlyres/6E783BC7-3467-484E-AD2A-29EF4A50432B/0/Mahone_2005_CS_White_Paper_SavingsEstimatingSavings.pdf&type=Excerpt&id=606909).

IV. Evaluation

A. Evaluation Standards

The primary standard in Demand Side Management evaluation is that evaluation must be conducted by an independent evaluator to guarantee the integrity of reported savings and to prevent conflict of interest. There has always been a separation of function between program implementers and program evaluators and a concern to place a wall between them.

In the early days of energy conservation programs (from approximately the end of 1970's through the early to middle to later 1980's), analysts on staff at utilities conducted many of the first evaluations. These first DSM evaluators were drawn from utility departments and work groups such as Forecasting, Load Research, and Rate Departments. These departments and work groups happened to have the required quantitative analytic skills when utility customer service departments initially started to create and implement the first energy conservation programs.

Separation at that time was already understood as an important principle. At the time separation was understood to require that the evaluation function would report up through a different high level manager or officer than the program implementation function so that the evaluators would be insulated from any appearance or actual conflict of interest to make programs look better than they were.

The standard in Demand Side Management evaluation is that evaluation must be conducted by an independent evaluator to guarantee the integrity of reported savings and to prevent conflict of interest.

An additional factor is that from the late 1980's through the first decade of the 2000's there was also a considerable and recurring problem with performance contractors who would propose that they deliver both the program and the evaluated savings results. When utility program contractor procurement practices did not catch this inherent error, this caused the introduction of unusual practices and disputes regarding inflated energy savings in several jurisdictions. The problem extended to various ways to game the appearance of impact results in utility customer information

systems.[14] When some part of the representation of energy use reduction in the customer information system is gamed, formal evaluation methods can fail even if applied by an independent evaluator. This is particularly the case when an unknown auxiliary variable is used to bias results.

In response to problems encountered and envisioned, and as DSM programs began to mature, regulators across jurisdictions required evaluation to be conducted by outside independent evaluators. By 1992, the requirement for independent evaluations was a general regulatory requirement.[15]

This separation of evaluation from program has always applied to program delivery agents managed by the DSM Administrator – an evaluation of its own activity by a program delivery agent under contract to the DSM Administrator is acceptable for the Delivery Agent's internal use. But an evaluation conducted by a Delivery Agent inherently appears as a conflict of interest and lacks regulatory value due to the problems of conflict of interest.

B. Transparency

Beyond this, DSM evaluation plans and reports are made public and are open to inspection by any party. Third party review of evaluations helps insure objectivity and promotes credibility, and, when necessary, revision of evaluation results.

C. Technical Resource Manual & Evaluation Guidelines

Most jurisdictions eventually develop a Technical Resource Manual (TRM) in which calculation methods, assumptions and (in some cases) default unitary values for energy savings and/or components used in the construction of energy savings and demand-reduction are curated and open for inspection by all parties. TRMs are generally considered a regulatory

[14] Peach, H. Gil, Pam Brandis, Michael J. Maranda, Ryan Miller, Anne West & Howard Reichmuth, "Evaluator as Fool, Reducing Uncertainty Through Evaluation", Brooklyn, New York: International Energy Program Evaluation Conference, August 2005. http://www.iepec.org/conf-docs/papers/2005PapersTOC/papers/031.pdf)

[15] A DSM Administrator will always have an evaluation manager and may have a few evaluation staff because it is necessary to maintain technical competence in this specialty area in order to choose evaluators and to manage evaluations.

document, structured with assistance from a consultant (and usually maintained by a consultant) with provision for periodic review by a coordinating committee representing parties. The values and calculating methods contained in a TRM are considered as reasonable until updated, usually as a result of evaluation results. Workable TRMs have specific procedures for periodically updating calculations and values.[16] Many jurisdictions also have a set of evaluation guidelines.[17]

D. The 2014 Evaluations

The evaluator carried out several evaluations for 2014 (Table 4).[18] As noted in the table, these evaluations are classified by type: process, market and impact. All seventeen program components or initiatives have one or more types of evaluation performed by the independent evaluator.

The primary focus in this Savings Verification study is on program impacts (GWh and MW).

[16] TRMs tend to incorporate material from earlier TRMs. For example, the Connecticut TRM is largely derivative from the Vermont TRM and the New York TRM began with the content of the Connecticut TRM. Similarly, Ohio's TRM consultant was a Vermont team and Indiana's TRM drew on both the Ohio and New York TRMs. California had one of the earliest database approaches similar to what we today call a TRM, but California is atypical in climate so many of its TRM values do not transfer well out of region and are used in other regions only as a last resort. In the Pacific Northwest, the Regional Technical Forum maintains a similar database of values, suitable for a more northern climate. NEEP maintains a Mid-Atlantic TRM.

[17] For example, the New York evaluation guidelines are at: http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf.

[18] This table reproduces the information of Table 1, Types of Evaluation Conducted by Program Component from the Econoler evaluation, Page 7. Econoler, Corporate Research Associates & Equilibrium Engineering, *2013 DSM Evaluation Reports; Efficiency Nova Scotia Corporation; Final Report*, April 20, 2014. Page numbers in the Savings Verification are referenced to the number systems within each report.

Table 4: Types of Evaluation.

Table 4: Types of Evaluation

	PROGRAM COMPONENT OR INITIATIVE	TYPE OF EVALUATION			
		Process	Market	Impact	
				Full-scale	Tracked Savings Validation
1	Appliance Retirement				X
2	Instant Savings	X	X	X	
3	Home Energy Assessment				X
4	Green Heat	X	X	X	
5	Low Income Homeowner				X
6	New Home Construction				X
7	Residential Direct Install	X	X	X	
8	Rental Properties and Condos Service				X
9	Solar				X
10	LED Holiday Light Exchange			X	
11	Home Energy Report	X	X	X	
12	Residential Financing	X	X		
13	Business Energy Rebates	X	X	X	
14	Custom Retrofit	X	X	X	
15	Custom New Construction			X	
16	Business Energy Solutions		X	X	
17	Codes and Standards			X	

Source: Econoler Evaluation of 2014 DSM Programs, Executive Summary, Table 1, Page 2.

V. Savings Verification Approach

The savings verification review was conducted as follows:

- During the year, we conducted a set of due diligence site visits for each program. These provide an independent view of each program and of any obvious problems. During the year, the site visits provide opportunity to call Efficiency Nova Scotia Corporation's attention to anything that needs to be explained or potentially fixed. It also keeps the Nova Scotia Utility and Review Board aware of any problems encountered.

- During the year we discussed use of the tracking system with ENSC staff. This provides an independent sense of how the tracking system is actually working.

- In review of the evaluations, we focused on the "installed" annual energy savings and demand reductions. [These are the annualized value of savings and demand reductions from the measures installed, regardless of what day of the year they were put into service. They do not represent the "realized" savings for the calendar year of the program, but are the savings that should be realized when each of the projects operates for a full year.] These numbers come from the tracking system and from the evaluator's adjustments.

- Generally, we did not check the evaluator's mathematical calculations, but did carefully check the evaluator's presentation of method for each program analysis. In particular, we reviewed the interaction, free ridership, spillover and net-to-gross approaches.

The Savings Verification study has one primary finding and several initial recommendations for the improvement of evaluation and programs. In addition, we endorse certain of the evaluator's recommendations that would improve measurement and evaluation.[19] Individual component findings and recommendations are also provided. Recommendations are summarized at the end of the report.

[19] Evaluator recommendations not explicitly endorsed are implicitly endorsed. Any exceptions are treated explicitly in the discussion on recommendations by program.

VI. General Findings

There is one primary finding:

1. **Savings Verification Finding No. 1: All but one of the 2014 Evaluations conducted by Econoler are within accepted industry frameworks.** Except for the Home Energy Reports evaluation, each evaluation was conducted within industry accepted evaluation frameworks and within the appropriate professional scope of seasoned Demand Side Management program evaluators. Beyond this, three of the evaluations can be classed as excellent because they systematically related evaluation steps to a recognized evaluation protocol.

   The single exception, Home Reports, was not conducted to industry standards because the evaluator was not provided the NSPI customer data to conduct random assignment to treatment and control groups, carry out data cleaning and selection and to conduct the primary evaluation analysis NSPI provided the customer data to the program delivery agent (Opower) instead of to the independent evaluator (Econoler).

   For Home Reports, data cleaning and preparation, the random assignment of customers to groups and the cross-sectional time-series panel study were carried out by the program vendor. *Within these constraints, the evaluator carried out the best secondary analysis possible but the analysis was not independent since the program vendor controlled the database and assignment of houses to treatment and control groups.*

   However, other than this one exception, each program evaluation is comprehensive: The structure and format of each impact evaluations is consistent and useful. Each has an Executive Summary that accurately reports the contents of the evaluation. Each provides key definitions of terms and acronyms up-front. Each has a good table of contents. Each has an explanatory methodological model diagram, followed by specification and discussion of methods employed in the evaluation. Citations throughout the evaluations to other studies are relevant. For each evaluation for which it applies there is an appendix with a careful and clear delineation of free ridership calculation methods and another appendix with an example of the impact calculation method.

General recommendations that apply throughout the evaluation effort are listed next, below. There are also individual Savings Verification recommendations listed in the program component or initiative sections that follow. Savings Verification recommendations are summarized at the end of the report.

VII. General Recommendations

1. **Savings Verification Recommendation No. 1:** The Savings Verification study recommends acceptance of the 2014 evaluation results for energy savings and for demand-reduction for all programs except Home Energy Reports.

The Savings Verification study recommends acceptance of the 2014 evaluation results for energy savings and for demand-reduction for all programs except Home Energy Reports.

2. **Savings Verification Recommendation No. 2: TRM.** Efficiency Nova Scotia should take steps to carefully establish a Nova Scotia TRM. This would facilitate both program planning and evaluation.

3. **Savings Verification Recommendation No. 3. Develop evaluation guidelines.** A separate document from a TRM is a set of evaluation guidelines covering methods, practice and reporting. The New York evaluation guidelines and appendices could be reviewed and modified to form an initial document.[20] This would set certain evaluation standards for future evaluations.

4. **Savings Verification Recommendation No. 4: Orient evaluations using accepted formal protocols**. As a general procedure, and as the evaluator carried out for three of the 2014 evaluations, for future evaluations, the evaluator should implement one of the formal evaluation protocols exemplifying "better practices". These protocols include the Uniform Methods Project evaluation protocols for typical DSM program types (UMP), the International Performance Measurement and Verification Protocols (IPMVP),

[20] New York State Department of Public Service, New York Evaluation Guidelines http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

ASHRAE Guideline 14, Measurement of Energy and Demand Savings, and the SEE Action energy efficiency program evaluation protocol guides. These protocols have been subject to extensive review across several jurisdictions and are coming into general use as a standard practice. They provide excellent structure and call attention to detailed considerations and specific steps required for evaluation of each type of program. The evaluator is, of course, always free to deviate from an evaluation protocol, but in the future implement evaluation using an industry standard protocol and should note chosen differences from the protocol and provide reasons.

This recommendation applies to the evaluator and to the DSM Administrator to specify as a requirement. The recommendation parallels and support several of the evaluator's recommendations for to the DSM Administrator for standardizing the DSM Administrator's tracking system records and calculation protocols.

5. **Savings Verification Recommendation No. 5: Savings Weighted Measure Lives.** The time dimension should be operationalized in the indicator system to supplement use of first year energy savings as a savings indicator. Programs differ in their time duration in providing kWh savings and demand reduction. Development of physical savings weighted measure lives for programs and program components and initiatives will provide a system for differentiating program efforts on this critical dimension and should be included in future impact evaluations. Keeping the first year energy savings as the sole performance criterion for assessing energy savings and demand reduction impacts makes standalone DSM behavioral programs appear equivalent to primarily technical and technical/regulatory DSM program. This mistake occurs as an artifact of the indicator system.

6. **Savings Verification Recommendation No. 6.** Several of the programs evaluated use the HOT 2000 model, which is eventually to be replaced by new software. The verification examination recommends the evaluator conduct a full evaluation including billing data model validation every three years, and again when the replacement software comes into use and the old software is retired.

VIII. Individual Program Component Review

Econoler conducted evaluations of seventeen DSM Administrator program components or initiatives (Table 4, Page 20). For the Savings Verification examination, we focus primarily on impact evaluation. These are designated in the two columns to the right of the Table 4 and are noted as either (1) full-scale evaluations or (2) tracked savings validations. Tracked savings validations are appropriate when a previous full-scale evaluation has established unitary savings values and when there has been no essential addition to the program measures or change in the nature of service delivery. These involve a review of the DSM Administrator's tracking system and of unitary savings values. These are checked for consistency with values from the previous evaluation and can be checked against values used in other similar jurisdictions. Sometimes this effort is supplemented with a files review (for program components for which paper files are maintained in addition to electronic tracking) and integration of interview and survey data from process and market evaluation to insure there has been no systematic change in the program.

A. Appliance Retirement Program (Tracked Savings)

The Appliance Retirement Program provides pick-up and environmentally sound disposal and recycling of components of household refrigerators, freezers and room air conditioners.

The evaluator used the Uniform Methods Project Refrigerator Recycling evaluation protocol for this evaluation[21]. There was some metering along with use of tracking information. The evaluation took into account free-ridership, secondary market impacts, induced consumption and internal spillover. Two institutional appliance replacement projects were also folded into this analysis. Electric interactive effects were taken into account.

In this evaluation, the embodied decisions on method and analytic approach are fully within the scope of performance by an independent evaluator. In addition, the evaluation is excellent because it follows a "better practice" universal evaluation protocol for this program type and because it incorporates some metering to provide grounding to calculated values. On the face of it, it may seem that an appliance retirement evaluation would be very simple. But it is not. For this evaluation, the evaluator carried out all the necessary detailed steps and included all factors that

[21] See: http://www1.eere.energy.gov/wip/pdfs/53827-7.pdf

might reduce program energy savings. Judgments and weightings were carefully carried out. *This is an excellent program evaluation*.

There are no evaluator recommendations for this program, which is a reflection of the maturity of the program in Nova Scotia and the performance of the program delivery agent, ARCA. Similarly, there are no savings verification recommendations for the evaluation of this program other than to continue future evaluations of this type with the same reference protocol, methods and careful focus evidenced here unless the evaluator finds some elements to improve.

B. Instant Savings Evaluation (Full-Scale)

Instant Savings is an "upstream discount" program component experienced by residential customers through discounts on energy efficient products when making purchases in stores. It is both a resource acquisition and a market transformation program. The evaluation deals with both dimensions, and emphasizes the need for program staff to stay in touch with the rapid evolution of markets for product types included in the program. Both program staff and evaluators are encountering rapid market changes that mean changing baselines.

The evaluation makes use of delivery agent and store staff interviews, in-store visits, intercept surveys and careful examination of unitary savings. This evaluation involves an individual detailed calculation for each type of instant savings product. A special focus is on changes in reference baselines, for example LEDs replacing CFLs. The evaluation requires borrowed data in several areas and uses relevant criteria for careful selection studies.

What would seem to be a simple evaluation of purchase of discounted energy efficient products actually has a high degree of complexity. This evaluation generally follows the Universal Methods Project Residential Lighting Evaluation Protocol.[22] In this evaluation, the included decisions on analytic approach, methods and quantitative values are fully within the scope of performance by an independent evaluator. In addition, the evaluation is excellent because it follows a "better practice" universal evaluation protocol for this program type. *This is an excellent program evaluation.*

[22] Universal Methods Project Residential Lighting Evaluation Protocol. See: (http://energy.gov/sites/prod/files/2015/02/f19/UMPChapter21-residential-lighting-evaluation-protocol.pdf).

The evaluator recommendations for this program are all process recommendations except two that directly affect calculations and classification. These are:

- Ensure that model number field is completed and accurate for all appliances.
- Ensure retailers enter all requested information.

The savings verification examination supports these impact-related evaluator recommendations. There are no savings verification recommendations for the evaluation of this program other than to continue future evaluations of this type with the same reference protocol, methods and careful focus evidenced here unless the evaluator finds some elements to improve.

The method, decisions and calculations performed by the evaluator are appropriate to this program type and within the scope of an experienced independent evaluator.

C. Home Energy Assessment (Tracked Savings)

Home Energy Assessment (HEA) provides financial incentives in the form of rebates or zero-interest financing to homeowners to reduce consumption of energy. It is focused on shell insulation measures to improve the building envelope in order to reduce the need for electric space heating. To measure performance, it incorporates a “test-in/test-out” residential audit including blower door testing.

The evaluation consists of a validation of the data in the program tracking system and inclusion of the Net to Gross Ratio based on survey results for free ridership and spillover. This year, the evaluator also included participants who started the program (received the “test-in” D Audit) but did not finish with the “test-out” E Audit. Results of a prior 2012 full-scale billing data model calibration analysis were incorporated in this study.

The incorporation of the “Unconverted D Study Savings” analysis of households that started the program, made improvement did not continue for the “test-out” E Audit is a constructive innovation in this year’s evaluation. These results, when added to results for households completing the program, strengthen the program impact. The demonstration of this effect suggests that past evaluations understated the actual savings for this program type.

The method, decisions and calculations performed by the evaluator are appropriate

to this program type and within the scope of an experienced independent evaluator. There are no evaluator recommendations for this program.

D. Green Heat (Full-Scale)

The Green Heat Program focuses on replacement and supplementation of heating systems by installation of equipment that uses fuel derived from renewable resources. Participants can choose between a rebate and a zero-interest financing option. The new equipment may either fully replace or supplement existing heating equipment. Incentives are available for central heating systems, as well as to supplement electrical heat with an area heater. The following categories of heating system are rebated:

- Wood and pellet stoves
- Wood and pellet boilers and furnaces
- Advanced pellet boilers and furnaces
- Geothermal heat pumps
- Central ducted heat pumps
- Air-to-water heat pumps

The evaluation was full-scale with process, market and impact evaluation. It included an interview with the program manager, five participant site visits, a survey of fifty-three participants and a unitary savings review. Unitary savings values were checked through literature review (or where comparison data was not available by developing engineering calculations). Free ridership was taken into account. The evaluator also followed up on the status of recommendations from 2012 and 2013.

The method and calculations performed by the evaluator are appropriate to this program type and within the scope of an experienced independent evaluator.

There are five formal evaluator recommendations to the DSM Administrator:

- **2014 GH-R1** – Consider using an initial system efficiency other than 100 percent when the air source heat pump or the geothermal heat pump is used as the initial primary heating system.

- **2014 GH-R2** – Improve the pre-screening process to ensure the use of exact information regarding the initial heating system type.

- **2014 GH-R3** -- The Evaluator recommends that ENSC use the updated tracked unitary savings value for central air source heat pumps. If air-to-water heat pumps account for a larger portion of Green Heat savings, ENSC should consider carrying out a calibrated model simulation of installed units or a metering study to measure and verify consumption and peak demand savings.

- **2014 GH-R4** -- The literature review did not show conclusive results regarding wood stoves and pellet stoves because no detailed metering studies on similar climates were identified. Additionally, since the tracked unitary savings values for these two types of equipment were updated in the 2013 evaluation using calibrated simulations, the Evaluator recommends that ENSC continue using the current tracked savings values.

- **2014 GH-R5** -- Similar to the results of the literature review for wood and pellet stoves, the Evaluator did not find any updated studies for wood and pellet furnaces/boilers. Since the current tracked unitary savings are based on simulations carried out by Dunsky Energy Consulting and Navigant Consulting, as well as a billing analysis conducted by Econoler in 2012, the Evaluator recommends keeping the current figures. Given the small number of projects installing these technologies through the program, it is hard to justify the need for conducting a metering study or a more thorough analysis.

Savings Verification Recommendation No. 7: Evaluator Recommendations for Green Heat. The savings verification examination supports each of the five evaluator recommendations for Green Heat.

In addition, since three of the five site visits found different original equipment than recorded in program records, this savings verification evaluation recommends that in the future this kind of disparate finding should immediately trigger a more extensive series of site visits by the evaluator.

Savings Verification Recommendation No. 8: Data Trigger. Should, in the future, the evaluator find an important pattern of discrepancies in small sample site visits, this situation should trigger an additional larger set of site visits to see if the pattern is more general.

E. Low Income Homeowner (Tracked Savings)

The Low Income Homeowner (LIH) Program is aimed primarily at building shell improvements. It includes building envelope measures, such as draft-proofing, as well as insulation for basements, crawl spaces, exposed floors, walls and attics, if required. The program also includes ventilation fans in bathrooms

and moisture reduction measures (such as covering dirt floors in basements) on a case-by-case basis. The program can replace old dehumidifiers, refrigerators and freezers with more energy efficient models.

The evaluation for 2014 reviewed tracking information parameters and included a complete re-examination of the EnerGuide point value as well as modelled energy consumption and a unitary savings review. The re-examination of the EnerGuide point value based on HOT 2000 data was competently carried out and resulted in a new EnerGuide point savings value (expressed in kWh per EnergyGuide point).[23]

The assumptions, methods and reasoning embedded in the evaluation are appropriate for this type of evaluation; the evaluator worked within the scope of an independent program evaluator.

F. New Home Construction (Tracked Savings)

The New Home Construction program, formerly Performance Plus, is designed to encourage homeowners and builders to exceed building code requirements for energy efficiency in new homes.

Measurement for this program is based on HOT2000 simulations. Once the house is built there is a final inspection and blower door test to verify the achieved efficiency level. The cost of the HOT2000 assessment is refunded for homes that achieve or exceed the Energy Star certification, EnerGuide 85 or higher, or receive Passive House certification. There is a possibility of an additional incentive of $1,500 for homes that exceed Energy Star certification and are expected to use less than 19,000 kWh based on HOT 2000 modelling.

The 2014 evaluation was based on validating tracked savings and re-evaluating the EnerGuide point value, free-ridership and spillover (market effects). The point value was developed from the Home Energy Assessment heating consumption data analysis (Home Energy Assessment also uses HOT 2000). A participant survey and a builder survey were used to revise free-ridership and spillover. The evaluation also contained a literature review.

This approach and the calculations employed by the evaluator are sound for this program type. Decisions embedded in the valuation are within the scope of an experienced independent evaluator.

[23] The coefficient of determination (R squared) is equal to 0.97 indicating a very strong relationship. With the current emphasis on effect size, this is a reasonable basis for acceptance of the result.

G. Residential Direct Install (Full-Scale)

The Residential Direct Install Program (RDI) provides direct installation of standard energy-efficient lighting and domestic hot water-heating products free of charge. In addition, participants may purchase specialty compact fluorescent lamps (CFLs) and/or LEDs at reduced costs. The products for 2014 were:

- CFLs (until July 1, 2014)
- LED lamps
- LED nightlights
- Faucet aerators
- Low-flow showerheads
- Pipe insulation
- Hot water tank wraps for electric water tanks
- Smart power controllers for AV equipment

The program serves both homeowners and renters. In addition a small Community Based Social Marketing pilot was included within this program.

For this evaluation, Econoler conducted a process and impact evaluation and a limited market evaluation. Interviews were conducted with the Program Manager and two Delivery Agents. There were fifty-three site visits, a quarterly telephone survey with (n=110) and an annual survey (n=500). Savings were calculated based on a unitary savings review.

The evaluation found that the switch to LEDs from CFLs and the addition of the smart power controllers resulted in increased energy savings. Although number of households served has declined from prior years, savings per household has increased. The approach and methods of calculation for this evaluation are sound. The embedded assumptions and decisions are also within the scope of an experienced an independent evaluator.

Data entry errors observed appear to inflate reported energy savings and this pattern of discrepancies was not recognized by the DSM Administrator's quality control review. Accordingly, Econoler provides the following recommendation:

RDI-R2 – Follow up with Delivery Agents to identify the cause of entry errors.

The Savings Verification study supports this recommendation by Econoler and, in additional requests that Econoler build a more extensive and systematic check on data entry errors for this program in the next evaluation.

Savings Verification Recommendation No. 9: Data Entry. The DSM Administrator should develop a focus on possible data entry errors for the Retail Direct Install program and the evaluator should include a more extensive and systematic check on possible data entry errors for Retail Direct Install in the next evaluation.

H. Rental Properties and Condos Service (Tracked Savings)

The Rentals Properties and Condos Service is the new name for the former Multi-Unit Residential Buildings program. It provides direct free-of-charge installation of energy-efficient products in both common space areas and rental units to landlords, renters or condominium owners. Installations in common areas are allocated to the Business Energy Solutions (BES) program, so the evaluation is confined to work completed in individual apartments or condominiums. Measures included for residential units for 2014 are:

- CFLs,
- LED lamps,
- Led Nightlights,
- Faucet aerators,
- Low-flow showerheads & shower wands,
- Pipe insulation and
- Hot water tank wraps.

The older 2013 evaluation was based on in-depth interviews with the Program Manager (PM) and the delivery agent (DA), as well as thirty-three site visits and a telephone survey of 45 participants.[24] The 2014 evaluation focused on a review of parameters and an update for gross and net savings along with a survey of nineteen landlords and property managers. The evaluation includes parameter validation, interactive effects and free ridership.

The evaluation approach and the assumptions embedded within calculations to determine evaluated savings are sound. The study and results are within the scope of an independent evaluation.

I. Solar (Tracked Savings)

The Solar Program includes both residential solar hot water and residential solar space heating (to offset part of the home heating needs met by existing traditional

[24] 2013 Multi-Unit Residential Buildings Evaluation, Page 8.

equipment). For 2015 there were thirty DHW systems and three air space heating systems. The DSM Administrator uses RETScreen analyses and Econoler applied adjustments established in the 2013 evaluation to establish gross and net savings.

This is a reasonable approach to evaluation of this program type for 2015. The checks and adjustments made by Econoler are within the scope of an independent evaluator.

J. LED Holiday Light Exchange (Full-Scale)

There were one-hundred and seventeen events for LED Holiday Light Exchange in 2014. The LED Holiday Light Exchange provides the opportunity to trade two sets of old inefficient holiday lights for one new set of LED lights. This is primarily a community outreach and awareness program.

The evaluation for the 2014 program is based on calculations using information in program records and an in-field survey, as well as a literature review to check results for consistency with results of similar programs elsewhere. Interaction effects were calculated and diversity was taken into account.

The assumptions and adjustments for this evaluation are reasonable. This is a good evaluation for this program type and methods and assumptions are within the scope of an independent evaluator.

K. Home Energy Report (Full-Scale)

The Home Energy Report analysis includes 2013 and 2014 since savings were not claimed in 2013. *This program differs from all other of the DSM Administrator's programs in that it is a standalone behavioral program*. In comparison, all other programs with impact evaluations for 2014 are centered on technologies or a combination of technologies and codes & standards. The central concept of the Home Reports program is to provide repetitive mailings to households with comparisons to undisclosed similar households in order to motivate behaviors that increase the energy efficiency of a household. These behaviors can include such things as taking shorter showers, lowering thermostats in Winter and raising them in Summer, choosing to participate in other technology based DSM programs, and purchasing behavior – for example, choosing to replace an appliance with a more energy efficient appliance rather than a less energy efficient selection in the market. In other words, for this program type, the internal physical mechanism that

generates the energy savings to be assessed is not well specified. Instead, from a physical perspective the savings (if any) are explained in a general sense as probably due to an aggregation over all participants and all non-participants (as two separate groups) of small changes in behaviors, participation in other DSM programs, and purchases (where, if it occurs, appliance purchases are defined as behaviors). The amorphous mechanism of the program is in contrast with primarily technical programs built around specification of particular equipment with known physical characteristics and technical/regulatory programs which specify a combination of technologies and codes and standards as the mechanisms for production of energy savings.[25]

The independent evaluator, Econoler, characterizes the Home Energy Reports program as composed of two program components, which are evaluated separately:

- The *Paper-Based* program component consists of the repetitive mailing of comparisons to similar households, plus suggestions or tips for saving energy.

- The *Web-Based* program component consists of participants also driven to an energy-efficiency website where they can become more fully engaged and also enter information that can trigger provision of more focused energy efficiency information and tips for saving energy.

1. The Paper-Based Component

In theory, evaluation of a Home Energy Report type program is conducted using a method of true experiment in which a group of homes is randomly allocated to either a *participant group* (the members of which will receive the repetitive comparisons to similar households) or a *control group* (the members of which will not receive the repetitive comparisons). If the experiment is conducted correctly according to protocol,[26] the difference in energy savings (and/or demand reduction) between the

[25] There are various social-psychological theories that might apply to motivating energy saving behaviors in a Home Reports program. The program type is amorphous as to the cause of energy savings in contrast to the other types of programs for which causation is crystal clear.

[26] State and Local Energy Efficiency Action Network, *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*, 2012. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. (http://behavioranalytics.lbl.gov)

two groups becomes the net impact of the program. Free ridership does not have to be taken into account due to the random assignment of households to groups (in theory, if random assignment is done correctly without any biasing factor included then since there is no differential selection to groups, they will both be assumed to contain the same proportion of free riders and in comparisons the free riders will cancel out). The evaluations can be conducted as pre/post studies or as post-only cross-sectional time-series panel studies. The panel study analysis method is a statistical "black box" method that works based on careful adherence to each step including assignment of houses to groups, data cleaning, and comparison. However, Figure 8 shows what happened.

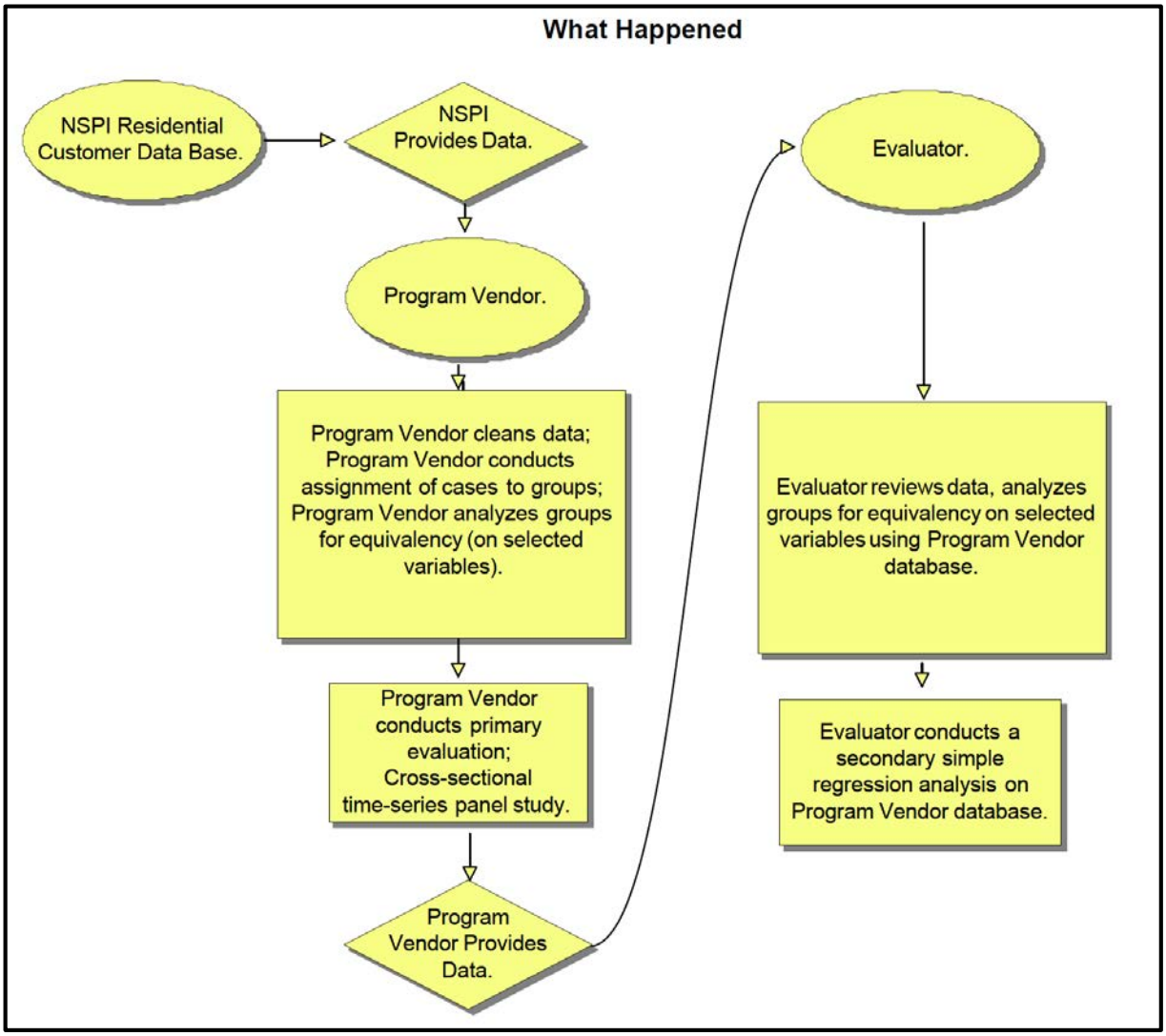


Figure 8: Actual Data Flow and Evaluation Performance.

What happened is that NSPI decided to provide the necessary customer data to the program vendor rather than to the independent evaluator. This created the data flow and sequence of events shown in Figure 8. *The problem is that keeping the independent evaluator out of the data stream until later in the project violates the scientific protocol for this evaluation.* No privacy interest is relevant here because the evaluator has to have the data to do the evaluation and eventually receives a database from the program vendor.

But handing the data off to the program vendor violates the principle of requiring an independent evaluation. It allows a program vendor to carry out the random assignment, clean and adjust the data set and carry out the primary evaluation. Figure 9 shows the appropriate data flow and evaluation performance.

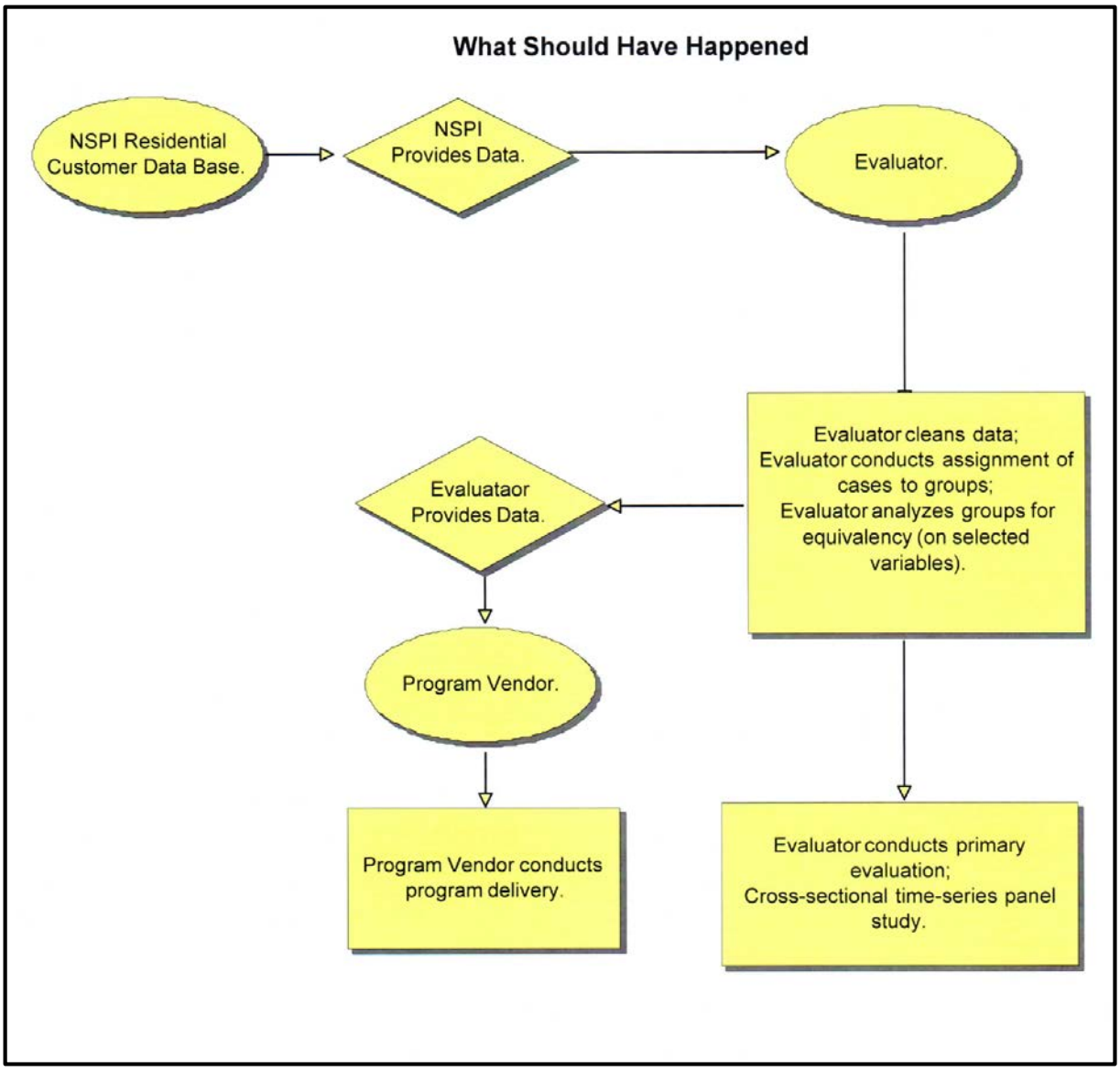


Figure 9: Required Data Flow and Evaluation Performance.

In Figure 9, NSPI would have properly provided the customer database to the evaluator. The evaluator would then have inspected the database, cleaned it and adjusted it. Then the independent evaluator would have carried out the random assignment to groups on which the integrity of the measurement system depends. The evaluator would then have handed off the customer database to the program vendor to conduct the program while the evaluator would have carried out the primary evaluation using a cross-sectional time-series panel study.

Because the evaluator was not permitted to control the database and conduct the random assignment and the primary analysis; and since these critical tasks were performed by the program vendor a critical conflict of interest was created. *The evaluator is then no longer an independent evaluator but a data analyst performing a weak secondary analysis and can only analyze data already adjusted, defined by, and provided by the program vendor.* By analogy, the legal equivalent is breaking the chain of custody of evidence in a criminal case – the evidence can no longer be admitted. Everything that flows from the point at which the program vendor controls the data and the assignment falls by analogy under the legal doctrine of fruit of the poisonous tree.

Evaluation of the Paper-Based component was conducted by the program vendor using a panel study. This is the primary evaluation (using a panel study post-only design) and the source of the evaluation results reported by the evaluator. The evaluator then conducted a weak secondary analysis using a simple regression model[27] limited by the data supplied by the program vendor. In addition, the evaluator replicated the program vendor's reported checks for equivalency of groups on selected variables (using the program vendor's datasets).

In conducting the evaluation, the evaluator reports that the program vendor says it developed two sets of participant and control groups. First, the evaluator reports the program vendor says they selected 20,000 NSPI residential customers and assigned them randomly to a participant treatment group of 10,000 and a control group of 10,000. Second, the evaluator reports that the program vendor says they selected NSPI's 110,000 highest residential energy users from NSPI residential customers and assigned them to a treatment group of 88,000 and a control group of 23,000.[28] The evaluator did not report the program vendor's separate results for

[27] Econoler Home Energy Report 2014 DSM Evaluation, Page 17.

[28] These numbers do not match exactly, they are off by 1,000 customers, but this does not affect the discussion.

each of the two sets of groups. The evaluator also did not report its analysis of the results for each of the two sets of groups.[29]

According to the evaluator, the program vendor asserts an annual average energy savings of 1.3% per home based on its panel study. The evaluator's analysis using a simpler method and the dataset provided by the program vendor did not disconfirm this result. Econoler's selected month comparison with Opower is shown in Table 5.[30].

Table 5: Opower & Econoler Savings Percentages for Paper Component.

Season & Month		Energy Savings Estimates	
		Opower	Econoler
Winter 2014	January	1.35%	1.53%
	February	1.40%	1.61%
	March	1.67%	1.86%
Summer 2014	July	1.78%	1.84%
	August	1.84%	1.88%
	September	1.88%	1.97%
Average for selected months		1.65%	1.78%

Econoler then applied the line loss factor determined total net savings at the generator as 34.498 (GWh) with 9.162 GWh of that in 2013 and 25.336 GWh in 2014.[31] Opower did not compute a demand reduction; however, Econoler independently computed the demand reduction as 3.938 MW at the generator overall with 1.046 MW in 2013 and 2.892 MW in 2014.

As noted above, a basic problem with the Paper Based analysis is that the primary evaluation for energy savings was conducted from start to finish by the program vendor rather than by the evaluator. As noted in the text box on Page 17, repeated here:

[29] No explanation is provided by Econoler for why Opower did not conduct a separate evaluation for each of these sets or why Econoler combined the two sets for a single analysis.

[30] Table 5 is based on Econoler Home Energy Report 2014 DSM Evaluation Table 10, Page 17.

[31] Econoler Home Energy Report 2014 DSM Evaluation, Page 18.

The standard in Demand Side Management evaluation is that evaluation must be conducted by an independent evaluator to guarantee the integrity of reported savings and to prevent conflict of interest.

This standard is specifically supported by the SEE Action evaluation protocol for behavior-based programs (Figure 10):[32]

3.2.1 Issue C: Avoiding Potential Conflicts of Interest

Evaluations of behavior-based efficiency programs should be managed in a way that produces the least potential for a conflict of interest to arise regarding the validity of savings estimates. Using a third-party evaluator that does not have a financial interest in neither the quantity of savings achieved nor the success or failure of the program (or its implementer or administrator) is the most transparent way to achieve this objective. This is particularly important if the evaluation being undertaken is intended to inform cost recovery or payment of incentives. In other situations (e.g., when an administrator is testing program marketing or design concepts or conducting a small pilot that involves technology demonstration and application), independent third-party evaluators may not be necessary to test preliminary program theories.

Table 3. Avoiding Potential Conflicts of Interest: Recommendation

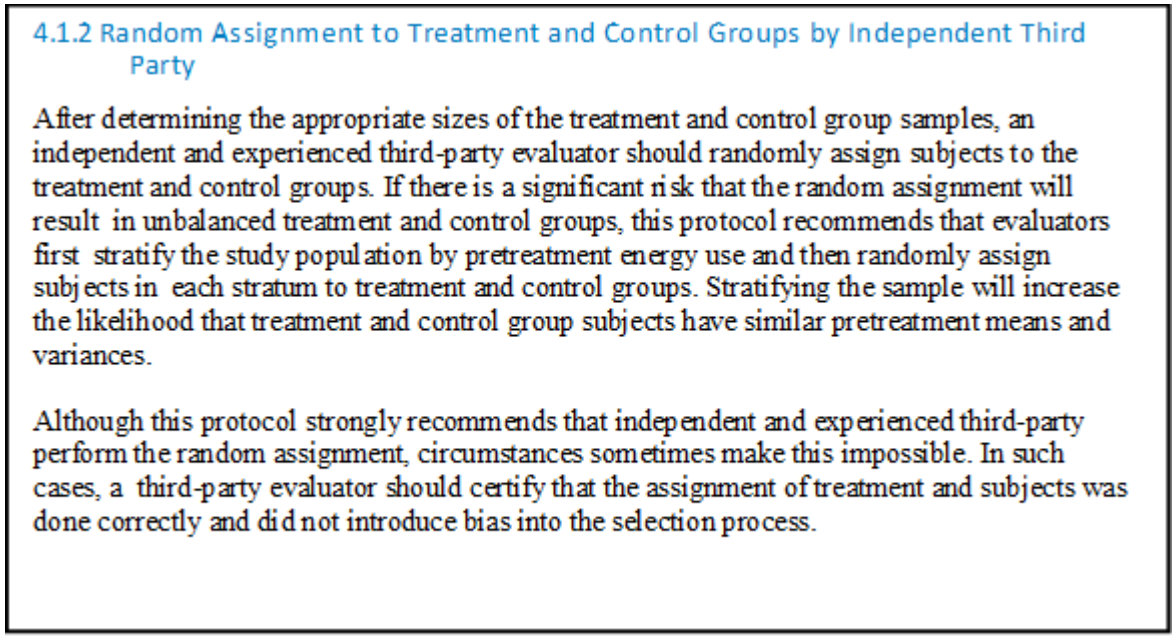
Star Rating	Condition
★★★★★	**An independent, third-party evaluator transparently defines and implements:** • Program evaluation • Assignment of households to control and treatment groups • Data selection and cleaning, including identification and treatment of missing values, outliers, and account closures, and the normalization of billing cycle days
★ **Not Advisable**	**Program implementer or sponsor** implements any of the above activities

Figure 10: SEE Action Evaluation Protocol.

The Universal Methods Protocol for behavior-based programs (Figure 11) also "strongly recommends" this standard *but also provides an alternative for situations in which independent evaluator assignment of households to groups is not possible*:[33]

[32] State and Local Energy Efficiency Action Network, *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*, 2012. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. (http://behavioranalytics.lbl.gov) See Section 3.2.1, Page 22.

[33] Universal Methods Project, Residential Behavior Protocol, August 2014. See http://energy.gov/sites/prod/files/2015/02/f19/UMPChapter17-residential-behavior.pdf



4.1.2 Random Assignment to Treatment and Control Groups by Independent Third Party

After determining the appropriate sizes of the treatment and control group samples, an independent and experienced third-party evaluator should randomly assign subjects to the treatment and control groups. If there is a significant risk that the random assignment will result in unbalanced treatment and control groups, this protocol recommends that evaluators first stratify the study population by pretreatment energy use and then randomly assign subjects in each stratum to treatment and control groups. Stratifying the sample will increase the likelihood that treatment and control group subjects have similar pretreatment means and variances.

Although this protocol strongly recommends that independent and experienced third-party perform the random assignment, circumstances sometimes make this impossible. In such cases, a third-party evaluator should certify that the assignment of treatment and subjects was done correctly and did not introduce bias into the selection process.

Figure 11: UMP Behavioral Protocol.

Econoler was not provided the dataset by NSPI,[34] so was prevented from meeting its scientific responsibilities for independent evaluation. As noted above, for the current evaluation Econoler carried out a simple regression as a check and replicated Opower's equivalency tests on a set of variables (to show the participant treatment group and the non-participant control group are essentially equivalent on these variables). Econoler, given this situation, has responsibly carried out the best, but inherently fatal flawed analysis, that could be done without access to the original database and without being able to conduct the random assignment. The analysis is not independent because it is constrained by use of the program vendor's datasets entailing whatever choices the program vendor made on data exclusion, data cleaning, assignment to groups and other matters.

Econoler proposes in recommendation **2014 HER-R3** that they conduct a billing analysis to compare the pre-participation electricity consumption of the treatment and control groups to determine if there was a pre-existing trend that might have influenced results. This does not fully solve the conflict of interest situation but it is a necessary step. Accordingly, the Savings Verification study provides the following recommendation:

[34] The industry standard of not permitting a program vendor to do the primary evaluation was well established when the project began. The SEE Action and UMP Behavioral protocols were not yet available.

Savings Verification Recommendation No. 10: Investigate Prior Trending. The independent evaluator should conduct an investigation to determine the existence or non-existence of an implicit trending in the participant treatment group or the comparison group or their combination prior to the random assignment of households to groups by Opower. The analysis should include five years of monthly back data.

It should also be noted that the existence of statistical significance for extremely small differences between group impact results for very large samples is a weak test because statistical significance is largely a function of sample size. In comparing large groups across many variables, many apparently meaningless relationships with very small effect sizes are statistically significant. And this is part of the problem with standalone behavioral programs: they require very large sample sizes to detect what turn out to be very small effects, such as the 1.3% net energy savings estimated by Opower here. And a little difference like that would only take a very light "thumb on the scale" to create.

- The tool of multiple comparisons can catch a structured difference between the participant treatment group and the non-participant control group *if and only if the right set of variables is tested*. For example, suppose the program vendor would have carried out the assignment of houses to groups using the occurrence of an e-mail address associated with the household or some other feature that remains unknown. That, in itself, would likely be enough to create a difference of 1.3% between the groups – because the existence of an Internet address would key to a complex of socioeconomic preponderancies. This is the problem of "unknown unknowns" and it is not solvable except by luck.

- In addition, the evaluator cannot confirm correct implementation of a randomization procedure to assign houses to group because they did not run this assignment and were not present when the program vendor ran the assignment.

- Also, since the evaluator did not have custody of the customer database at this point they cannot determine what alterations were made to the database or if a biasing factor was introduced within the reported random assignment.

These facts lead to the second Savings Verification Finding and to Savings Verification Recommendation No. 11:

Savings Verification Finding No. 2: No independent evaluation. There is no independent evaluation for the Home Energy Reports program because the evaluator was initially kept from the data. The database cleaning and adjustment and the random assignment were controlled by the program vendor. The evaluator could only report on what the program vendor asserted that it did with the dataset and the assignment of houses to groups. The analysis conducted by the evaluator cannot be characterized as independent.

- *The results of this study are not adequate to support the energy savings and demand reduction calculations primarily because the program vendor conducted the primary evaluation.* The vendor's evaluation is identical to the vendor's savings claim.

- The DSM Administrator did not have a copy of the customer database, so could not track and question the actions and assertions of the program vendor with regard to the data and assignment to groups. This removed the DSM Administrator and the data tracking system as a checking mechanism. The DSM Administrator's numbers are the program vendor's claims.

- The evaluator's secondary and weaker analysis did not disconfirm the program vendor's evaluation results, but the evaluator's analysis is not independent because it was constrained by the program vendor's database and choices.

Savings Verification Recommendation No. 11. Run additional comparisons. The evaluator should run additional comparisons using a NSPI database rather than the program vendor's database, including prevalence of e-mail address by group to see if by luck a systematic difference between the participant treatment group and the non-participant control group can be detected. This kind of check will require a copy of the original database from NSPI in its original form before being provided to the program vendor.

Because the claimed energy savings are running at 1.3% and this is in the range of similar electricity Home Energy Reports studies elsewhere from the approximately four vendors who offer the program, the NSUARB could consider acceptance of the results with qualifications even though the scientific protocol was broken and the results are fatally flawed. This would be an administrative solution based on the claim of savings in reference to claimed percentage savings from similar projects and not based on the analysis conducted by the program vendor or the evaluator. This possibility is indicated Savings Recommendation No. 12, below.

Savings Verification Recommendation No. 12: Possible acceptance with conditions. The Program Vendor's self-evaluated impact results for Home Energy Reports could be tentatively accepted on the basis of being within a reasonable range, subject to the evaluator's continued investigation of prior trending and additional equivalence checks by the independent evaluator.

Finally, given the problems with the standalone behavioral program, we might ask the question, "is any of this worth the problems?" We return to this overarching question in a summary for Home Energy Reports, following discussion of the Web Based component.

2. The Web-Based Component

Evaluation of the Web-Based component was independently conducted by Econoler and was unable to establish any significant level of savings for this component. Econoler selected energy efficiency tips with reasonable participation that were not covered by other DSM programs and constructed unitary savings values for each tip. The selected tips are:

- Energy-efficient dishwasher,
- Energy-efficient room air-conditioner,
- Insulate windows with heat-shrink film,
- Shave a minute off shower time,
- Turn off the computer at night,
- Wash clothes with cold water, and
- Hang laundry to dry

For each tip a unitary savings value was constructed. Then in a single survey of 340 participants in the Opower treatment group who were driven to the website, the participants were asked about relevant behaviors over the past twelve months. For example, survey results showed sixteen participants each purchase an energy-efficient air conditioner instead of a non-efficient air conditioner and from this kind of data an implementation ratio was developed.[35] Similar implementation ratios were developed for each tip. For this analysis, free ridership was developed from the survey (since the control group was not included in the analysis). Net energy

[35] Econoler Home Energy Report 2014 DSM Evaluation, Pages 20-21.

savings were computed as gross energy savings times the net-to-gross ratio. There was high free ridership because in answering the survey, many who responded said they would have taken the action anyway, without receiving the in Home Energy Reports. Econoler was not able to establish any significant level of savings associated with the Web Based component.

This analysis is weak in that there is a single survey to request information on baseline and final conditions. The means there is recall bias, social desirability bias, and leading questions bias occurring in the survey situation. Given the attrition table there is also response bias in reaching the small subset of participants who volunteer to complete the survey. This leads to a recommendation for future work of this kind:

Savings Verification Recommendation No. 13: Baseline vs. Program Surveys. For future examinations of Web Based program components, consider surveys at different points in time so that baseline is established at least a year prior to the program effects survey.

3. Summary

Home Energy Report programs, introduced in 2008, have attracted participation by many utilities. Though initially pilots, some utilities now consider them DSM programs. The Home Energy Reports program type of evaluation has the aura of scientific evaluation because it is based on random allocation of households to treatment and control groups. The numerical savings claim (1.3%) by the program vendor is within the range of other Home Energy Reports projects.

However, having the program vendor evaluate its own program is not compliant with good procurement practice or long-standing DSM standards. Whatever the scientific aura, the acceptance of a program vendor's evaluation as the program evaluation of impact is not proper due to inherent conflict of interest. The evaluator did a creditable job of functioning under difficult conditions. However, the evaluator's work is secondary to the evaluation conducted by the program vendor and is not independent (because the evaluator used the program vendor's database and allocation and reported the program vendor's evaluation results).

As shown in the Relative Value charts of Figures 6 & 7, standalone behavioral programs have a much lower savings value than a first year savings analysis would suggest. Standalone behavioral programs can be valuable in an emergency

situation that requires civic cooperation to yield immediate energy savings. But, under normal conditions, they are not valuable as DSM programs because they have little value in resource planning and are squishy. This leads to a final recommendation for this program type – to declassify it as a DSM program since it has no real resource planning value under normal conditions.[36] However, behavioral approaches have a positive role as enabling strategies in support of technical DSM programs.[37]

Savings Verification Recommendation No. 14: Enabling Strategy. The NSUARB should consider ruling out standalone DSM programs as useful except in times of energy emergency and require that behavioral approaches be classed along with marketing, promotion, finance and information as enabling strategies, not as DSM programs.

L. Residential Financing (No Impact Evaluation)

This is a zero interest buy-down program with sixty-six participants in 2014. Program savings are allocated to other programs (HEA, Solar and Green Heat). The financing enables all necessary work to be done as a single expenditure.

The evaluation is process evaluation based on in-depth interviews with the Program Manager, the financial partner and with ten participants. The evaluation found that most participants would have postponed the work without the financing component, and that the financing program was pivotal.

[36] While the TRC test is not a part of the evaluation, there is also a question regarding proper costing of customer dollars for this program type. And, since the physical mechanisms of causation of energy savings are amorphous it is not clear that actions taken are all cost effective for the customer. For example, the appliance program is limited to certain (Tier 3) models but this program is not. And, at root a question is whether resource planning needs ephemeral behavioral savings in 2014 or can rely on it in planning in the same way that primarily technical programs can be relied upon. These question would have to be answered by the resource planning function but are noted here as relevant.

[37] For treatment of a range of successful enabling strategies, see: Fuller, Merrian C., Cathy Kunkel, Mark Zimring, Ian Hoffman, Katie Lindgren Soroye and Charles Goldman, *Driving Demand for Home Energy Improvements: Motivating Residential Customers to invest in Comprehensive Upgrades that eliminate Energy Waste, avoid High Bills, and spur the Economy*. Berkeley, Californian: Lawrence Berkeley National Laboratory, September 2010. (http://emp.lbl.gov/sites/all/files/REPORT%20low%20res%20bnl-3960e.pdf)

This is a carefully focused and successful process evaluation. The evaluator has produced relevant and useful results.

M. Business Energy Rebates (Full-Scale)

This program provides prescriptive rebates to business, non-profit and institutional organizations (abbreviated "BNI"). Originally, Business Energy Rebates (BER) was a mail-in rebate program. It now includes a second participation approach by applying instant rebates through electrical distributors.

The prescriptive measures for mail-in rebates are grouped into several categories: agricultural, commercial lighting, commercial refrigeration, compressed air, HVAC, motors and variable speed drives, commercial laundry, commercial kitchen, commercial solar thermal heating, and commercial water heating.

The instant rebates provided through electrical distributors are classed within the following product types: CEE electronic ballasts, CEE T8 lamps, high bay induction lamp fixtures, LED decorative walkways, LED downlights, LED docklights, LED garage canopy lighting, LED high bay fixtures, LED troffer, LED outdoor wall pack fixtures, LED strip lighting, T5 high bay fixtures, T8 high bay fixtures, lighting occupancy sensors and LED tubes for fluorescent lighting. Specific products are subject to change.

This evaluation includes in-depth interviews with the Program Manager and eight distributors as well as fifty-eight on-site visits, a survey of twenty-seven mail-in participants and four quarterly surveys of sixty-eight mail-in participants. The evaluator also followed up on the status of nine open 2012 and seven 2013 evaluation recommendations. Econoler provides seventeen recommendations to the DSM Administrator. The evaluation corrects for impact accounting problems and recommendations include changes to be made by the DSM Administrator so that correction in the evaluation is not necessary.

Impact is developed from a review of unitary savings values used in tracking. Econoler evaluated parameters used in the tracking savings calculations by conducting a literature review of technical reference manuals and evaluation studies. Where information was not available, Econoler developed engineering calculations. Free ridership and interactive effects were included.

The evaluation of the Business Energy Rebates program is clearly written and direct, well-reasoned and thorough. This evaluation is within the scope of an experienced

independent evaluator.

N. Custom Retrofit (Full-Scale)

This Custom Retrofit program for business, non-profit and institutional (BNI) customers provides an incentive to customers to implement retrofit projects that are deemed to produce significant annual savings and peak load demand reduction. The program provides both technical and financial support. Technical support for these projects typically begins with scoping (feasibility studies). Incentives are based on simple cost payback periods and are tailored to meet the situation of individual customers. There were 100 Custom Retrofit projects in 2014, in three Tracks (there is an "in-between" Track for compressed air projects).[38] Track 1 projects must save between 5,000 and 20,000 kWh of electricity per year. Track 2 eligible projects must save at least 20,000 kWh. Technological solutions are developed specially for each individual project. Although the program saves energy and reduced demand, the focus is primarily on energy savings.

Two other component programs, Energy Management Information Systems (EMIS) and Existing Building Commissioning (EBC) are included within the Custom Retrofit evaluation. EMIS is targeted to medium and large industrial customers and it provides industrial customers with funding required to audit their systems, develop an EMIS application plan and offset the cost of installing metering and monitoring equipment. Verified EMIS savings must be derived from consistently monitored improvements within the facility.[39] EBC implements no-cost and low-cost improvements in the context of building retrocommissioning for institutional and commercial customers.

The evaluation contains an impact, process and market evaluation. Determination of project savings relies primarily on the DSM Administrator's tracking database with adjustment factors determined via participant interviews and on-site visits. Free ridership and spillover are determined. The diversity factor and interaction factors are already accounted in the tracking database.[40] The 2014 evaluation was based on an in-depth interview with the Program Manager (PM), 22 on-site visits, a telephone survey of 11 participants and a review of 27 project files (22 Custom Retrofit, 3 EBC and 2 EMIS).[41] The evaluation also included a review of the DSM

[38] 2014 Econoler Custom Retrofit Evaluation, Figure 3, Page 11.

[39] 2014 Econoler Custom Retrofit Evaluation, Page 29.

[40] 2014 Econoler Custom Retrofit Evaluation, Page 24.

[41] 2014 Econoler Custom Retrofit Evaluation, Page vii and Figure 1, Page 3.

Administrator's response to five open recommendations from the 2012 evaluation and five recommendations from the 2013 evaluation. Econoler provided nine recommendations in the 2014 evaluation.

This Savings Verification Study specifically supports two of Econoler's recommendations:

Econoler Recommendation 5: Require EMIS service providers to track and present savings using the IPMVP.

Econoler Recommendation 6: Plan on conducting a thorough impact evaluation of EMIS and EBC in 2016.

Savings Verification Recommendation No. 15: IPMVP. Require EMIS service providers to track and present savings using IPMVP and conduct a full-scale impact evaluation of EMIS and EBC in 2016.

In addition the evaluator should focus on electrical measurement.

Savings Verification Recommendation No. 16: Address Electrical Measurement. For future evaluations of this program, the evaluator should provide a table to explain the frequency and type of direct electrical measurement involved in the projects, how the measurement was performed and which entities carried out the measurement. The evaluator should provide a partition of energy savings and demand reduction claimed according to the percentage based on direct electrical measurement and the percentage developed without electrical measurement.

This evaluation is within the scope of an experienced independent evaluator.

O. Custom New Construction (Full-Scale)

This Business, Non-Profit and Institutional (BNI) program provides prescriptive rebates to organizations for the implementation of above-code building envelope, lighting and HVAC design implementation for commercial new construction projects. There were thirteen projects in 2014. The program has two pathways, (1) energy modeling and (2) prescriptive.

The 2014 evaluation is based on an in-depth interview with the Program Manager, nine site visits, full technical documentation reviews and a billing

calibration.[42] The evaluator used ASHRAE Guideline 14-2002 and IPMVP. Free ridership was not applied due to the selection a new baseline method – the baseline is the original design intent for the project.[43]

The evaluator provided five recommendations. This Savings Verification Evaluation supports two of the five recommendations:

Savings Verification Recommendation No. 17. Calibration for Custom New Construction. Structure project monitoring to gather more data to conduct a conclusive calibration billing analysis in 2015 and develop the diversity factor used in peak demand savings by building it up from an individual measure basis.

In place of the evaluator's recommendation to use NECB 2011 as a baseline, the Savings Verification Evaluation recommends examining what is happening in Nova Scotia with regard to enforcement in practice of the new code. Generally there is a lag of some years between the updating of a code and either strict code enforcement or the gradual drift of practice to conform to the new code. Code enforcement is usually not well funded and there are typically not enough inspectors. Also training on the new code provisions may not be timely or complete. Code inspectors usually have too much work to focus on entire code, and unless there is special intervention by funding additional positions, placing DSM Administrator staff in the code enforcement agency and special training, the energy provisions of the code are not typically seen as a high priority. Because they are overburdened and can't do everything, Inspectors tend to focus on health and safety code provisions. In keeping with a focus on physical reality and "what's real" the following recommendation is provided:

Savings Verification Recommendation No. 18. Use current practice for Custom New Construction. The DSM Administrator should move to the new, more stringent, baseline construction designs of NECB 2011 when it finds them in current practice, but not before. We recommend code not be adopted as baseline until it is proven to be in common practice.

The evaluation is competent and well expressed. The evaluation is within the scope of an experienced and professional independent evaluator. This is an excellent evaluation, though limited by available data.

[42] 2014 Econoler Custom New Construction Evaluation, Page vi.

[43] 2014 Econoler Custom New Construction Evaluation, Page vi. The baseline selected takes the place of free ridership; if free ridership were calculated with this baseline it would be double counting.

P. Business Energy Solutions (Full-Scale)

Business Energy Solutions (BES) provides turnkey implementation, beginning with a free on-site energy assessment for small businesses. There is a substantial "buy-down" of full installation cost and many kinds of energy-efficient measures are included in the program. Service for common areas of the Rental Properties and Condos (RP&C) program are included in BES. For 2014, the program also included a LED Blitz pilot for facilities such as shopping malls that use more than 350,000 kWh annually. There were 781 regular program participants plus 51 RP&C common area participants (representing 105 separate building) and 42 additional participants in the LED Blitz pilot for a total of 854 participants.

Measures include:

- Efficient lighting measures
- Occupancy sensors and lighting controls
- Ductless mini split heat pumps
- Electronically commutated motors (ECMs)
- Outside air economizers
- Hot water heat pumps
- Hot water tank wraps
- Pipe insulation
- Spray valves
- Faucet aerators
- Vending machine misers

The evaluator conducted five site visits along with the Delivery Agents and analysis of unitary savings. Free ridership and internal spillover were estimated using the participant annual and quarterly surveys. Interaction effects are taken into account. The evaluation was divided into four categories: BES lighting measures, BES non-lighting measures, RP&C common areas and the LED Blitz. The evaluator provided nine recommendations for 2014. The Savings Verification evaluation supports all nine evaluator recommendations.

Savings Verification Recommendation No. 19: Business Energy Solutions. Implement all nine of the Econoler recommendations for the Business Energy Solutions program.

In addition, we recommend that the problems noted in the "Barriers" section of the evaluation be systematically addressed by the DSM Administrator.

Savings Verification Recommendation No. 20: Business Energy Solutions Barriers. The DSM Administrator should systematically address the problems noted under "Barriers to Participation".[44]

This evaluation is sound and is within the scope of an experienced professional independent evaluator.

Q. Codes and Standards (Full-Scale)

Energy savings and demand reduction impacts of Codes and Standards are evaluated Econoler evaluated but are accounted outside the DSM Administrator's DSM portfolio.

This evaluation develops the energy savings and demand reduction impacts of new regulations for:

- Electric motors,
- General service incandescent reflector lamps,
- Dry-type transformers,
- Self-contained commercial refrigerators and freezers,
- Large air conditioners and heat pumps,
- Single package vertical air-conditioners and heat pumps,
- Standby power,
- External power supplies,
- Incandescent Lamps (75W and 100W),
- LED street lighting, and
- New energy codes for houses.

Most of these improvements in minimum energy performance standards (MEPS) were enacted at the federal level and are implemented through NRCan as part of Canada's Clean Air Regulatory Agenda. The improvements reduce energy use as well as greenhouse gas and air pollution. The minimum standard for LED street lighting and the implementation of new energy codes for housing and small buildings are Nova Scotia initiatives. Since the code for residential boilers has not been implemented, it is not included. And digital television adapters are not included due to lack of data.

For 2014, the primary sources of information for analysis were NRCan and industry

[44] Econoler Business Energy Solutions, Page 12-14.

trade associations. Each analysis requires data and understanding of the workings and context of a specific market (market characterization of products with quantity by size, type, efficiency level, etc.).[45] The evaluation estimates impact in terms of energy and demand savings, taking into account both shipments to Nova Scotia and non-compliance rates. The evaluation is based on national analysis results maintained by Natural Resources Canada, adjusting by ratio using Nova Scotia population and work with secondary data sources (including in-depth interviews with trade associations).

This evaluation type requires a considerable amount of research in order to understand products and markets. The evaluation is within the scope of an experienced and professional independent evaluator.

Because the regulatory requirements for improved technology come from the federal level or the provincial level, the DSM Administrator does not take DSM credit for energy savings and demand reduction caused in this program area. However, the DSM administrator does work to support the implementation of federal and provincial standards and this requires both some continuing staff involvement in the area of standards and also a continuing evaluation effort.

A key question on codes and standards is the definition of the administrative level for code enforcement. Often building codes are enforced at the county level of government and there can be wide variation in building practice and code enforcement by county. This is an area in which technology training and code enforcement pilots can be opportunities for an efficiency utility to boost compliance with federal and provincial codes and standards. Such pilots allow assignment of incremental savings (increased compliance levels) to the utility. The MEEA Midwest Regional Codes Conference is one source for different utility approaches to codes and standards. (http://www.mwalliance.org/policy/midwest-regional-energy-codes-conference). There are also models in Washington State and in California for utility codes and standards work. One of the first steps in advancing this work is to reach an agreement among parties on assigning some credit for codes work to the efficiency utility. Like market transformation programs, codes and standards may require five or more years of up-front work before savings results begin.

[45] Econoler, Introduction of New Codes and Standards in Nova Scotia, 2014 DSM Evaluation, Page xi.

IX. Summary of Findings & Recommendations

Savings Verification Finding No. 1: All but one of the 2014 Evaluations conducted by Econoler are within accepted industry frameworks. Except for the Home Energy Reports evaluation, each evaluation was conducted within industry accepted evaluation frameworks and within the appropriate professional scope of seasoned Demand Side Management program evaluators. Beyond this, three of the evaluations can be classed as excellent because they systematically related evaluation steps to a recognized evaluation protocol. (Page 22)

Savings Verification Finding No. 2: No independent evaluation. There is no independent evaluation for the Home Energy Reports program because the evaluator was initially kept from the data and the database and random assignment were controlled by the program vendor.
(Page 42)

Savings Verification Recommendation No. 1: The Savings Verification study recommends acceptance of the 2014 evaluation results for energy savings and for demand-reduction for all programs except Home Energy Reports. (Page 23)

Savings Verification Recommendation No. 2: TRM. Efficiency Nova Scotia should take steps to carefully establish a Nova Scotia TRM. This would facilitate both program planning and evaluation. (Page 23)

Savings Verification Recommendation No. 3. Develop evaluation guidelines. A separate document from a TRM is a set of evaluation guidelines covering methods, practice and reporting. The New York evaluation guidelines and appendices could be reviewed and modified to form an initial document.[46] This would set certain evaluation standards for future evaluations. (Page 23)

Savings Verification Recommendation No. 4: Orient evaluations using accepted formal protocols. As a general procedure, and as the evaluator carried out for three of the 2014 evaluations, for future evaluations, the evaluator should implement one of the formal evaluation protocols exemplifying “better practices”. These protocols include the Uniform Methods Project evaluation protocols for typical DSM program types (UMP), the International Performance Measurement and Verification Protocols (IPMVP), ASHRAE Guideline 14, Measurement of Energy and

[46] New York State Department of Public Service, New York Evaluation Guidelines http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf

Demand Savings, and the SEE Action energy efficiency program evaluation protocol guides. These protocols have been subject to extensive review across several jurisdictions and are coming into general use as a standard practice. They provide excellent structure and call attention to detailed considerations and specific steps required for evaluation of each type of program. The evaluator is, of course, always free to deviate from an evaluation protocol, but in the future implement evaluation using an industry standard protocol and should note chosen differences from the protocol and provide reasons. (Pages 23-24)

Savings Verification Recommendation No. 5: Savings Weighted Measure Lives. The time dimension should be operationalized in the indicator system to supplement use of first year energy savings as a savings indicator. Programs differ in their time duration in providing kWh savings and demand reduction. Development of physical savings weighted measure lives for programs and program components and initiatives will provide a system for differentiating program efforts on this critical dimension and should be included in future impact evaluations. Keeping the first year energy savings as the sole performance criterion for assessing energy savings and demand reduction impacts makes standalone DSM behavioral programs appear equivalent to primarily technical and technical/regulatory DSM program. This mistake occurs as an artifact of the indicator system. (Page 24)

Savings Verification Recommendation No. 6. Several of the programs evaluated use the HOT 2000 model, which is eventually to be replaced by new software. The verification examination recommends the evaluator conduct a full evaluation including billing data model validation every three years, and again when the replacement software comes into use and the old software is retired (Page 24)

Savings Verification Recommendation No. 7: Evaluator Recommendations for Green Heat. The savings verification examination supports each of the five evaluator recommendations for Green Heat. (Page 29)

Savings Verification Recommendation No. 8: Data Trigger. Should, in the future, the evaluator find an important pattern of discrepancies in small sample site visits, this situation should trigger an additional larger set of site visits to see if the pattern is more general. (Page 29)

Savings Verification Recommendation No. 9: Data Entry. The DSM Administrator should develop a focus on possible data entry errors for the Retail Direct Install program and the evaluator should include a more extensive and systematic check on possible data entry errors for Retail Direct Install in the next evaluation. (Page 32)

Savings Verification Recommendation No. 10: Investigate Prior Trending. The independent evaluator should conduct an investigation to determine the existence or non-existence of an implicit trending in the participant treatment group or the comparison group or their combination prior to the random assignment of households to groups by Opower. The analysis should include five years of monthly back data. (Page 41)

Savings Verification Recommendation No. 11. Run additional comparisons. The evaluator should run additional comparisons, including prevalence of e-mail address by group to see if by luck a systematic difference between the participant treatment group and the non-participant control group can be detected. (Page 42)

Savings Verification Recommendation No. 12: Possible acceptance with conditions. The Program Vendor's self-evaluated impact results for Home Energy Reports could be tentatively accepted pending results of an investigation of prior trending and additional equivalence checks by the independent evaluator. (Page 43)

Savings Verification Recommendation No. 13: Baseline vs. Program Surveys. For future examinations of Web Based program components, consider surveys at different points in time so that baseline is established at least a year prior to the program effects survey. (Page 44)

Savings Verification Recommendation No. 14: Enabling Strategy. The NSUARB should rule out standalone DSM programs as useful except in times of energy emergency and require that behavioral approaches be classed along with marketing, promotion, finance and information as enabling strategies, not as DSM programs. (Page 45)

Savings Verification Recommendation No. 15: IPMVP. Require EMIS service providers to track and present savings using IPMVP and conduct a full-scale impact evaluation of EMIS and EBC in 2016. (Page 48)

Savings Verification Recommendation No. 16: Address Electrical Measurement. For future evaluations of this program, the evaluator should provide a table to explain the frequency and type of direct electrical measurement involved in the projects, how the measurement was performed and which entities carried out the measurement. The evaluator should provide a partition of energy savings and demand reduction claimed according to the percentage based on direct electrical measurement and the percentage developed without electrical measurement. (Page

48)

Savings Verification Recommendation No. 17. Calibration for Custom New Construction. Structure project monitoring to gather more data to conduct a conclusive calibration billing analysis in 2015 and develop the diversity factor used in peak demand savings by building it up from an individual measure basis. (Page 49)

Savings Verification Recommendation No. 18. Use current practice for Custom New Construction. The DSM Administrator should move to the new, more stringent, baseline construction designs of NECB 2011 when it finds them in current practice, but not before. We recommend code not be adopted as baseline until it is proven to be in common practice. (Page 49)

Savings Verification Recommendation No. 19: Business Energy Solutions. Implement all nine of the Econoler recommendations for the Business Energy Solutions program. (Page 50)

Savings Verification Recommendation No. 20: Business Energy Solutions Barriers. The DSM Administrator should systematically address the problems noted under “Barriers to Participation”. (Page 51)

X. References

Donaldson, Stewart L., Christina A. Christie & Melvin M. Mark, eds., *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* Los Angeles, London, New Delhi, Singapore: Sage Publications, 2009.

Dunsky Energy Consulting, "Enabling Strategies: Overview of 5 Regions' Budget Ratios", January 2, 2014.

Econoler, Corporate Research Associates & Equilibrium Engineering, *2014 DSM Evaluation Reports; Efficiency Nova Scotia Corporation*.

Gellings, Clark W. & John H. Chamberlin, *Demand-Side Management Planning Concepts & Methods*, Second Edition, Liburn Georgia: The Fairmont Press, 1992.

Gellings, Clark W. & John H. Chamberlin, *Demand-Side Management Planning*, Liburn, Georgia: The Fairmont Press, 1993.

Heberlein, Thomas A., *Navigating Environmental Attitudes*. Oxford & New York: Oxford University Press, 2012.

Keating, Kenneth M., Ruth L. Love, Terry V. Oliver, H. Gil Peach & Cynthia B. Flynn, "The Hood River Project; Take a Walk on the Applied Side," Pp. 112-118, *The Rural Sociologist*, Vol.5, No. 2, May 1985.

Khawaja, M. Sami and James Stewart, "Long-Run Savings and Cost-Effectiveness of Home Energy Report Programs," Cadmus, 2014. (http://www.cadmusgroup.com/papers-reports/long-run-savings-cost-effectiveness-home-energy-report-programs/).

Mahone, Douglas, Nick Hall, Lori Megdal, Ken Keating & Rick Ridge, "Codes and Standards White Paper on Methods for Estimating Savings," Prepared for Marian Brown, SCE in support of Statewide NRNC MA&E, April 7, 2015. (http://www.cpuc.ca.gov/NR/rdonlyres/6E783BC7-3467-484E-AD2A-29EF4A50432B/0/Mahone_2005_CS_White_Paper_SavingsEstimatingSavings.pdf&type=Excerpt&id=606909)

Morse, William L. & H. Gil Peach, "Control Concepts in Conservation Supply," *Energy*, Vol. 14, No. 11, Pp. 727-735, 1989

Peach, H. Gil, Pam Brandis, Michael J. Maranda, Ryan Miller, Anne West & Howard Reichmuth, "Evaluator as Fool, Reducing Uncertainty Through Evaluation", Brooklyn, New York: International Energy Program Evaluation Conference, August 2005. (http://www.iepec.org/conf-docs/papers/2005PapersTOC/papers/031.pdf)
State and Local Energy Efficiency Action Network, *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*, 2012. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. (http://behavioranalytics.lbl.gov)

State and Local Energy Efficiency Action Network, *Energy Efficiency Program Impact Evaluation Guide, 2012*. Prepared by Steven R. Schiller, Schiller Consulting, Inc., www.seeaction.energy.gov.

New York State Department of Public Service, New York Evaluation Guidelines http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/NY_Eval_Guidance_Aug_2013.pdf.

The MEEA Midwest Regional Codes Conference, papers and reports. See: (http://www.mwalliance.org/policy/midwest-regional-energy-codes-conference).

Universal Methods Project Residential Lighting Evaluation Protocol. See: (http://energy.gov/sites/prod/files/2015/02/f19/UMPChapter21-residential-lighting-evaluation-protocol.pdf).

[1]Universal Methods Project, Residential Behavior Protocol, August 2014. See http://energy.gov/sites/prod/files/2015/02/f19/UMPChapter17-residential-behavior.pdf

Westling, Hans, *Cooperative Procurement: Market Acceptance for Innovative Energy-Efficient Technologies*, Nutek, Stockholm, Sweden, 1996, 141 pp.

.

Suggested Citation:

Peach, H. Gil & John Mitchell, *Verification Review of Program Year 2014 Evaluation Results/Report for the Nova Scotia Utilities and Review Board.* H. Gil Peach & Associates/Scan America®, March 2015.