

Measuring Program Performance From Cradle to Grave

Marc Collins and Andrew Bishop, EM&V, Ontario Power Authority

ABSTRACT

A frustration for many demand-side management (DSM) program managers is that evaluation results and feedback are not available until after a program is out-of-market. The organization may eventually know how well the program did, but only when it is too late to make early- or mid-course process corrections and improvements. This paper will show how one organization has addressed that problem in a number of innovative ways for several of its key DSM initiatives.

The first method involved using continuous awareness and participation polling while programs are in market as an “early warning system” and then incorporating those results into a separate formal program evaluation. A case study for a program which made discount coupons available to the public at retailers for electricity-efficient products will be used to demonstrate the survey and EM&V process and results.

The second method involved very early process evaluation feedback for an HVAC upgrade program that involved two separate out-sourced delivery agents—one for managing field contractors and the other for rebate processing. The feedback allowed helpful program modifications to occur and be proven even before the final evaluation report was issued.

The third opportunity to aid program management decisions is possible due to a new approach to evaluating demand response (DR) initiatives. As part of the development of an innovative DR evaluation framework, the Ontario Power Authority (OPA) is moving away from a focus on after-the-fact, which often means after the peak weather season, assessment of DR performance, to an *ex ante* approach. In other words, the most relevant way to value DR resources is on the basis of their future potential or insurance value. It is this forward-looking assessment of DR initiatives that should form the basis of their cost-effectiveness calculation and contribution to electricity system resources. By looking forward, program managers now have the information they need to make in-market, in-season adjustments. In addition, the data collection required to perform the evaluation analysis also provides a useful database for program management scenario and administration purposes.

Introduction

The OPA is an independent corporation reporting to the Ontario Minister of Energy & Infrastructure. Its mandate is to ensure an adequate, long-term supply of electricity for Ontario. Its objectives include engaging in activities in support of the goal of ensuring adequate, reliable and secure electricity supply and resources in Ontario and that promote electricity conservation and the efficient use of electricity and also to collect and provide information relating to medium and long term electricity needs of Ontario. In support of those objectives the OPA funds a number of electricity efficiency and demand response initiatives. Government or utility-sponsored DSM activity in Ontario had almost ceased for more than decade prior to the OPA’s creation in 2005.

Example 1 – Integration of Market Research

The EKC Power Savings event encourages Ontarians to conserve electricity by providing in-store discount coupons at participating major retailers, including major hardware, home & auto and discount department chains, for certain electricity-efficient products at point-of-purchase. The products

covered by the program vary from season to season and may as an example contain ENERGY STAR®-qualified compact fluorescent floodlights, T8 fluorescent light fixtures, ENERGY STAR®-qualified light fixtures, heavy-duty timers for pools, hot tubs, and furnace filters. The program ran in two flights: Spring (April 18 – June 15) and Fall (October 1 – November 16).

During the program period, market research was done with eligible Ontarians to monitor awareness levels and other metrics as a means of providing interim feedback to program and marketing managers and stakeholders inside and outside the OPA on program and advertising performance, profiling participants of each program for purposes of targeting future programs and communications and learning the impact of having the program in market on awareness and image of the Every Kilowatt Counts brand.

Polling was done with a representative sample of 200 Ontarians every week beginning with benchmark research prior to launch date and continuing for eight months, past the end date of the final program in market—a total of over 4,000 completed interviews. Reporting was done weekly at a macro level and at scheduled points throughout the program periods.

Data collection was done by an independent market research company with their proprietary in-house online panel. The survey instrument asked about awareness of the program (by name and description) and recognition of its advertising, current and intended participation. It also collected a “net promoter score” (a measure of satisfaction and loyalty) for the program. Finally, the questionnaire asked awareness of and attitudes towards the Every Kilowatt Counts brand as an entity.

A representative sample of the population was used because of the broad nature of the Power Savings Event’s target group. At the same time, because much of the program promotion used media such as newspaper and in-store point-of-purchase materials it meant that one of the objectives (effect of program marketing on the awareness and image of EKC) relied on feedback from the general public rather than only its target audience.

Online data collection was used for two reasons. First, it facilitated the presentation of the program name and description, featured product images and coupon images to participants to optimize the recall accuracy. Second, it enabled exposing the communications to participants: newspaper and point-of-purchase communications were displayed as images. Traditionally, one criticism of online panels has been sample bias. That risk was weighed against the sample bias inherent in low and still-declining response rates to landline telephone surveys and the increased quality of response in online surveys, and the decision to go the online route was clear.

Weekly data collection was done to allow measurement before, during and after the program’s promotion period to map program and advertising awareness to media plans and to measure the retention in awareness once program promotion ended, as input to future communications and media planning. As well, to-date results, updated daily, were made available to the Ontario Power Authority as an early-warning system for use in “course correction” should program or advertising awareness be below acceptable levels.

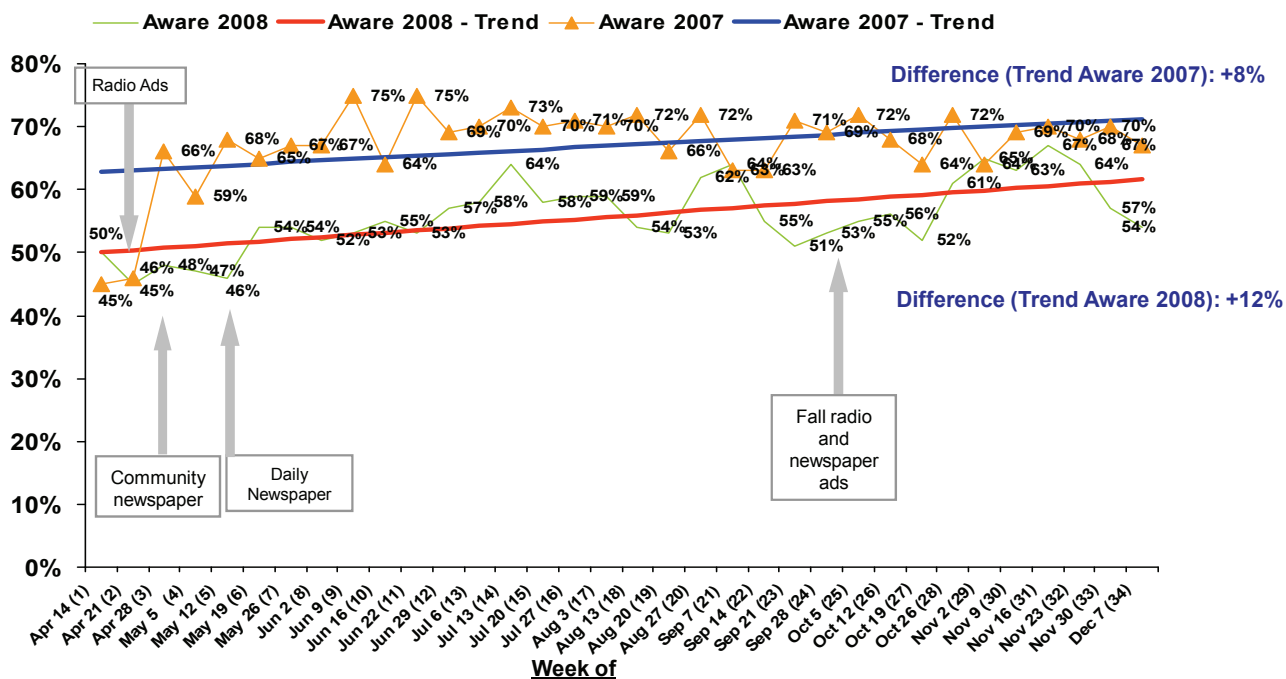
Both program and advertising awareness were substantially lower during the Spring flight than those seen during the same period in 2007 (see Figure 1).

Figure 1



Awareness of Every Kilowatt Counts Power Savings Event

Awareness of the Power Savings Event has increased over time, peaking during the Fall campaign. Overall, about one-half of Ontarians can recall hearing about the Power Savings Event.



A2. The "Every Kilowatt Counts Power Savings Event" encourages Ontarians to conserve electricity by providing coupons for electricity-efficient products. Ontarians can participate in the program by purchasing and redeeming coupons for ENERGY STAR®-qualified compact fluorescent floodlights, T8 fluorescent light fixtures, ENERGY STAR®-qualified light fixtures, heavy-duty timers for pools, hot tubs and spas, and furnace filters. Are you aware of the "Every Kilowatt Counts Power Savings Event" by its name or description? 2007 Base: 7443 2008 Base: All respondents n=6576

Participant profiling showed there was no specific demographic group, segment or region that accounted for this drop in awareness. Market research was done shortly after the Spring flight ended to identify opportunities to increase the effectiveness of program promotion materials to be used in the Fall flight. Focus groups were done with the target group, whose feedback led to a more focussed, single-minded creative communications approach—one which stood out at shelf level (see Figures 2 and 3).

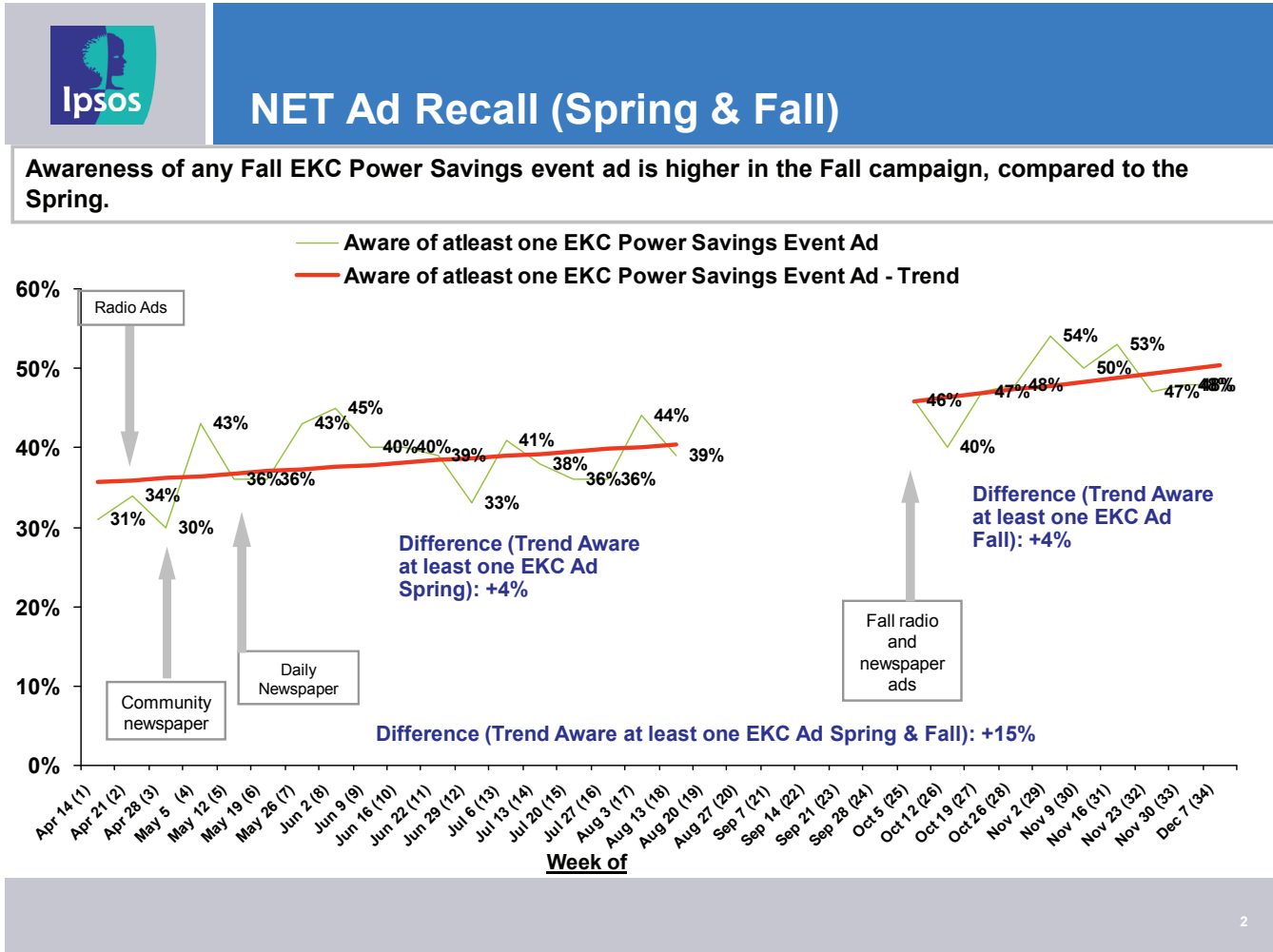
Figure 2

SPRING

FALL

Directional awareness was much higher in Fall than in Spring, so the creative strategy which was informed by the focus groups formed the foundation for the 2009 Spring campaign. In short, the mid-program market research did the job it was meant to do: it identified an opportunity area in time for changes to be made which contributed to the overall success of the 2008 EKC Power Savings Event.

Figure 3



EM&V Participant Consultation and Research

A key input to impact and process reviews is participant feedback. On the impact side, participant feedback is critical in developing program specific free rider or other net-to-gross ratios. On the process side, participant feedback is critical in understanding how effectively a program was delivered to market, how successful various marketing and awareness components of the program encouraged participation and in facilitating the development of lessons learned. These program design critiques can be applied to future year iterations of the same program or other programs that target a similar audience.

However, due to the retailer-based approach taken in this initiative, the lack of a program/participant tracking system for EKC presented a significant challenge to evaluators. Participants in the program were unknown to the OPA—or any other agency or participating retailer for that matter—making follow-up survey work difficult. This meant unique and creative approaches were needed to facilitate participant consultation.

Through the prior evaluation effort of the 2007 EKC program, evaluators utilized a cold-calling approach to reach and consult with participants. To enable this process, a call-centre was established that randomly dialled tens of thousands of Ontarians to first identify and then, where appropriate, question the

participants on their experiences with the program. Although this methodology was successful in achieving its goal – in large part due to the significantly high rates of participation across the province (over 1,000,000 coupons were redeemed)—the process took significant time and money to complete.

The significant level of participation in the 2007 EKC program was due in part to the effective marketing and awareness efforts of the program, as well as the recognizable EKC brand associated with the initiative. It was also due to the measures included in the program. CFL's, as well as other more universally accepted measures were incented through both 2007 campaigns and the take-up on these products was considerable.

However, for the 2008 campaigns, the measures included purposely did not have the same levels of market penetration or awareness. In an attempt to push technologies that were further down the market transformation curve, less popular products were selected. It was anticipated that the levels of participation seen in the 2007 campaigns would not be repeated in 2008. Without these same levels of participation, a cold calling approach to participant consultation would be even more time consuming and expensive to complete.

The prompt provision of participant feedback is considered critical for two reasons. First, it ensures that any identified program issues or potential improvements can be addressed quickly—prior to the subsequent iterations of the initiative. This is especially important in the case of the EKC where a significant level of co-ordination between the OPA and participating retailers is required. This co-ordination ensures that products promoted through the EKC program are available at retailer locations during the program period. Any recommended change to the retailer component based on participant feedback would take months to implement.

Second, consulting with participants as soon after participation as possible ensures that the most accurate information to inform both process and impact evaluations is gathered. Through the 2007 evaluation effort, participants for both Spring and Fall programs were being interviewed several months following the end of the Fall campaign. This meant Spring EKC participants were being consulted up to 6 months following their experience with the program. This lag time between participation and interview resulted in a significant reduction in the quality and accuracy of information gained through participant consultation.

A combination of the new, “less popular” measures offered in the program and the expected high cost and lengthy time required to collect participant feedback necessitated a revised approach to the collection of participant feedback. Working in combination with the OPA's Market Research team, it was agreed that the use of an independent market research company's proprietary in-house online panel was a worthwhile option to investigate. The rationale for this exploration included the following: the costs would be finite; we could potentially collect information after each campaign; and the information collected could be provided to the program managers in a timely manner to inform potential program design changes (in time for design changes for the Fall campaign, based on Spring results).

EM&V undertook its first test of the panel approach following the Spring 2008 EKC campaign. During the delivery of the program, the evaluation contractor and market research firms whose panel would be used to conduct the research were selected (using competitive procurement processes). Following, a team consisting of staff from the OPA's Market Research and EM&V groups, the Evaluation Contractor and the Market Research Firm worked together to develop the survey instrument that would be used to collect information from 2008 EKC spring campaign participants. OPA EKC program managers were also consulted and provided detail on additional elements related to the program they would like to see investigated through the effort. Components of this instrument would carry through to the surveys used in the program evaluation many months later.

While the program was in market, the final survey instrument was developed and the Market Research Firm readied the instrument for its panel members. One week following the completion of the Spring campaign, the survey was sent to panel members for response.

Utilizing the panel approach, the OPA was successful in obtaining the necessary participant information to inform both the impact and process evaluations. Of the six measures included in the campaign and investigated through the survey, all but one obtained the desired sample of participant responses. However, the results for all measures were considered significant and conclusive by the evaluation contractor (although higher error bands surrounding the result of the one measure that did not obtain the desired sample were developed). Participant feedback was also obtained promptly following the delivery of the program – all participant feedback was obtained within five weeks of the program conclusion – ensuring the most reliable and accurate data was collected.

Significant cost savings were achieved through the use of the on-line panel. Compared to the cold-calling effort, the on-line approach cost approximately one-fifth the amount. Thus, the exploration into the use of online panels was successful. Based on the success of the Spring experiment, the methodology was repeated for the Fall Campaign evaluation with similar success. Of course, in addition to satisfying EM&V quality and cost concerns, the program managers/operators also benefitted from timely feedback that allowed them to implement changes at least one program cycle sooner than waiting for the formal program evaluation results.

Example 2 – Early Process Evaluation Feedback

The OPA is not a utility, though we are responsible for DSM programs like the ones a utility would deliver. There is a network of local distribution companies (LDC's) that cover the entire provincial territory and often they are the lead or a partner in delivering programming. In other cases, usually where splintered delivery across 75 LDC's would result in diseconomies of scale, the OPA works with a third-party delivery agent to manage and execute a program. In the case of this program, which was a new entry to the marketplace after many, many years of no DSM programming in Ontario, there was also a dual program manager split. This tri-partite arrangement, combined with pressure to produce energy and demand savings quickly, caused some teething pains. We knew that early process evaluation feedback could assist all parties in resolving administrative and delivery issues that were already present in the early stages of the program roll-out.

The 2007 Hot Savings Rebate and Cool Savings Rebate Programs were province-wide energy efficiency initiatives that targeted the nearly 2.2 million residential electricity consumers in Ontario with existing central space heating and space cooling systems. Originally launched in 2006, the 2007 programs were delivered through heating, ventilation and air conditioning (HVAC) contractors across the province. The program itself was managed by the OPA, which used an HVAC industry professional institute for promotion within the contractor community and traditional program administration and a separate firm provided rebate processing and incentive fulfilment.

Any HVAC contractor in Ontario had the opportunity to participate in the program, at no cost, upon successful completion of the mandatory program training module developed for the program. The training was developed to inform contractors how to:

- 1) Appropriately administer the program (including required participant documentation and process for incentive fulfilment);
- 2) Market the benefits of energy efficient HVAC equipment to potential participants.

The original version of the Cool Savings Rebate Program ran from April to September 2006, and was designed to encourage homeowners to:

- Tune-up their air-conditioning system prior to the cooling season (July and August);
- Replace their existing central cooling system with an Energy Star[®] qualified air conditioning system (which exceeds current regulatory efficiency standards); and
- Purchase, install and have programmed (by a participating contractor) a programmable thermostat.

The Hot Savings Rebate Program ran from October 2006 to March 2007, and focused on promoting:

- The replacement of existing lower efficiency furnaces with higher-efficiency models equipped with an electronically commutated motor (ECM);
- The replacement of existing lower efficiency central air conditioning systems with a higher efficient Energy Star[®] system (which exceeds current regulatory efficiency standards); and
- Purchase, install and have programmed (by a participating contractor) a programmable thermostat.

A secondary benefit to the Hot Savings Rebate Program was the environmentally friendly manner in which old mercury based thermostats are decommissioned.

The 2007 Cool Savings Rebate Program ran from April 2007 – March 2008 and built upon the objectives of the previous programs, to save energy and reduce demand in the residential sector by:

- Increasing the number of participants installing high or mid-efficiency furnaces equipped with an ECM (when replacing existing lower efficiency equipment);
- Increasing the number of contractor programmed programmable thermostats installed to control space heating and cooling;
- Encouraging homeowners to replace existing lower efficiency space cooling equipment with an Energy Star[®] qualified cooling system; and
- Encouraging homeowners to have their cooling system maintained on a regular basis prior to the start of the cooling season.

EM&V Research/Evaluation Contribution

It became clear early in the process of designing an evaluation plan that there were operational issues that had the potential to threaten potential savings by delays, confusion, misrepresentation or all three. EM&V staff committed to the OPA program manager to devise a series of interviews that would probe the issues. In an environment of heightened tension and anxiety that existed amongst the three delivery parties, the process evaluation effort turned-out to be like advice from a “neutral third party.” It allowed all three parties to voice their concerns to a literal third party, which was helpful in a situation where roles and responsibilities were still evolving amongst the three entities. The overall context was still very much one of start-up mode, with a prevalence of lack of formal program documentation and procedures.

Exception handling for rejected rebate forms (that HVAC contractors failed to fill-out thoroughly or correctly) was a classic example of where applying diligence by the fulfilment house could cause embarrassment to the industry institute. Another by-product of the somewhat unusual tri-partite delivery accountability arrangement involved differences of perspective regarding the auditing and verifying of the installations themselves. The institute wanted to believe that its members were diligent and honest. The fulfilment house wanted confidence that it was paying-out rebates only for the right amount for legitimate work. The OPA wanted both. Delays in the payment of rebates were perceived as excessive “leniency” by one party, but were likely the result of under-resourcing and inability to pay more attention by parties that had multiple accountabilities.

This program design and initial delivery “state” included a lack of clarity in a number of areas to an extent that it directly led to relationship friction amongst the three delivery partners. Roles were evolving as the pressure to deliver results led to expediency-based decision-making that could not afford to respect the initial program design, where there were specifics, or perceived roles.

The process evaluation helped to defuse this potential powder-keg by allowing all parties to focus on the results as the primary motivator. Each could express their views of how the results were potentially compromised by the actions or inactions of others, without having the instantaneous (and probably negative) feedback that occurs in regular meetings/interactions. The evaluation opened the lines of communication between delivery partners, helped to build a positive and mutually respectful group dynamic and a shared appreciation for the unique issues faced by the three separate program delivery partners.

Changes or improvements that resulted from the process evaluation findings and recommendations include the following:

In addition to enhancing the day-to-day experiences of each delivery partner, this improved communication also alleviated the vast majority of the issues that existed in the program’s delivery. As suggested by the evaluator, a necessary first step to developing a well-functioning team was to collaboratively work to clarify the accountabilities of each partner. Through this process each party became fully aware of their own “roles and responsibilities” relating to the following:

1. How program marketing (e.g., point-of-purchase) and other materials are created, reviewed and approved and who is responsible for their distribution to participating contractors;
2. Who is responsible for managing, maintaining and updating the Contractor enrolment and training component;
3. The proper process for marketing and advertising campaign development/management – including the approvals processes necessary;
4. Clarification on the accountabilities of each party with regard to all rebate submission, processing, recording and payment processes;
5. The correct approach to responding to general program enquiries from both contractors and participants – including, depending on who has made the enquiry, who needs to be informed of the enquiry, who is involved with crafting responses and who officially corresponds with the party;
6. Discussions around rebate exception handling – how is this properly managed? Who validates the exceptions? How are contractors that frequently submit rebate forms which are unclear or incomplete handled?; and
7. Scheduled group discussions to review the current delivery of the program and, where appropriate, propose refinements to improve the ongoing operation of the initiative.

Additionally, regular meetings (monthly at a minimum) were also suggested by the evaluator. These meetings, as well as the roles and responsibilities clarification, help ensure that an ongoing healthy relationship exists between the three program delivery partners and that any issues which present themselves are dealt with efficiently and effectively.

A formal program audit process – a recommendation of the evaluator – was collaboratively developed by the program partners. During the 2008/09 program evaluation, the program partners designed an audit and inspection process that reflected the suggestions of the evaluator.

The design of this process ensures that the OPA:

1. Retains control over the auditing of installed measures and will work with program partners to address any anomalies that are observed;

2. Has the ability to perform frequent spot investigations of rebate applications that raise ‘red flags’ by the rebate processor; and
3. Ensures that as anomalies are addressed, necessary adjustments to processes and procedures are made, recorded and communicated to all involved parties.

The OPA has also adjusted the rules of rebate eligibility to ensure that participants agree to a home visit (in principal) in order to qualify for their rebate. This inclusion helps eliminate any issue that may have hindered the value of home-visits conducted for this evaluation—namely, that anyone knowingly engaged in fraudulent activity would be highly unlikely to agree to a home verification visit.

Through the newly introduced audit process, verification of between 1% - 3% of installed measures is targeted for the initial year, with a potential adjustment up or down depending on findings.

Although the audit process is owned and managed by the OPA, all delivery partners are fully engaged in the effort. This engagement included audit design input as well as regular attendance at all project update and reporting meetings. This audit initiative is a solid example of the level of integration, coordination and co-operation that now exists between partners.

Example 3 – *Ex Ante* Demand Response Approach

The third example of evaluation directly assisting program operations at the OPA involves the demand response (DR) suite of programs. The OPA developed a suite of DR initiatives in response to ever-increasing electricity system capacity restrictions. The overall political savings targets for Ontario were, in fact, expressed in capacity terms—6,300 MW by 2025, representing approximately 20% of projected capacity. Remembering the dearth of pre-existing DSM activity in Ontario, the OPA had a significant challenge to ramp-up DR results in a short time frame. The evaluation approach discussed here will illustrate how DR program managers can more effectively target their product(s), squeeze efficiency from them and balance individual DR program streams across the DR portfolio.

Note: technical details regarding the methodologies employed in the OPA’s approach to DR evaluation can be found in the following paper, also presented at IEPEC 2009: “A Comprehensive Framework for Evaluating Demand Response in a Resource Planning Context,” by Nicole Hopper, Ontario Power Authority, Toronto; Dr. Stephen George and Josh Bode, Freeman, Sullivan & Co., San Francisco.

The OPA looks at DR from a resource planning perspective. This is partly because resource planning is a central part of the mandate of the organization, but also because the aforementioned system capacity issues pushed us toward initiatives that had potential to be reliable enough to be treated as resources by system planners and system operators. EM&V therefore wanted to assess and value the true peak/extreme weather benefits and costs of DR. We needed to do this with very little time to roll-out and test programs or go through “normal” program cycles.

The OPA had already developed a framework for evaluating energy efficiency resources based on best practices from decades of efficiency program evaluation in other jurisdictions. However, for DR, evaluation methods are less well established and there are fewer examples to borrow from. It would be necessary to develop forward-looking estimates of the expected performance and value of DR resource options. An examination of practices in other jurisdictions revealed that no existing framework would directly suit the OPA’s needs, though it was determined that California offered the most relevant case study for Ontario.

The California Public Utilities Commission (CPUC) has adopted a set of DR load impact estimation protocols (CPUC 2008) and has made progress toward development of a DR cost-effectiveness framework; both take a forward-looking approach. In 2008, the OPA hired FSC, who had contributed to the California protocols, to develop a DR evaluation framework for Ontario. The resulting framework borrows from the

California process and products, but incorporates lessons learned from their implementation and includes some changes to suit the Ontario market.

The Ontario context is that the majority of consumers in the Province pay fixed retail rates for electricity. Large industrial users who are “market participants” are an exception. These facilities pay wholesale market prices for their electric commodity. A Smart Meter initiative is underway that, when complete, will result in all classes of customers having interval meters installed. A pilot time-of-use (TOU) rate is in effect for some of the first customers to receive Smart Meters, and the Ontario Energy Board (OEB) is overseeing development of a TOU rate structure that would eventually (by 2011) apply to all retail customers across the Province.

Currently, the OPA offers three DR programs to non-residential customers: “DR-1,” a voluntary demand-bidding program; “DR-2,” a permanent load shifting program; and “DR-3,” a capacity program. In addition, the OPA manages “*peaksaver*^{®1},” a residential and small-commercial direct load control program that is operated by local distribution companies (LDCs). LDC’s may become more involved in delivering DR soon. The Ontario IESO also maintains a resource of large electricity users that can be called to provide relief during system emergencies.

Given this backdrop, the OPA required the DR evaluation framework to:

- be defensible, transparent, and appropriately value the specific attributes of DR;
- be relevant throughout the resource planning and program delivery cycle;
- fully recognize the option (or insurance) value of DR;
- specify minimum requirements to enable DR resources from a potentially large number of sources to be included in resource plans on an apples-to-apples basis, without imposing burdensome evaluation requirements; and
- be relevant in a shifting policy and regulatory environment, and for a wide range of potential DR resource options.

Lastly, there is the issue of public reporting. OPA wanted to focus on what impacts should be estimated and what should be reported in order to facilitate comparisons between DR resources and to allow trained reviewers to assess the accuracy, precision, and robustness of the evaluation methodology.

There are eight load impact protocols that respect differences in DR resource-type characteristics. They also detail the day types and event conditions that are used as the basis for the analysis—some *ex post* and some *ex ante*.

In addition, a cost-effectiveness framework, centred on the Total Resource Cost (TRC) Test is used. In order to calculate the TRC, it is necessary to determine the avoided cost of the capacity that DR replaces. Since DR is a peaking resource, there is general agreement that the avoided cost should be based on the avoided cost of a single-cycle combustion turbine (SCGT). However, not all DR resources are available in all hours of the year, and many have restrictions on their timing and number of hours of use that do not apply to an SCGT. For example, air conditioner (A/C) cycling programs provide larger load reduction resources as temperatures increase and may not be available at all during the winter months. Ontario-specific options include using the relative loss-of-load probability (LOLP) and relative expected unserved energy (EUE).

A relatively accessible, spreadsheet-based tool accompanies the framework—for cost-effectiveness and predictive, or *ex ante* projections of load impacts. This is the key to how EM&V is helping program managers. Instead of waiting for *ex post* evaluation results an entire program cycle or more after-the-fact, program managers can now do their own projections with a high degree of accuracy.

¹ ® Trademark of Toronto Hydro Corporation. Used under license.

They can also predict how well the program will perform from a cost-effectiveness perspective under various weather-year scenarios.

In the case of one program, the correlation between the difference in contracted maximum load and the contractual baseline calculation was extremely close to 1. This means that program managers can track subscribed customers/load and then predict impact results, with confidence, “live.” Knowing in advance what types of loads, in what time frames (and if Ontario used locational pricing, where) produce which impacts and what those impacts are worth, allows for huge improvements in targeting marketing and sales. In the current economic downturn, it permits a sophisticated assessment of how much cost-effective DR could occur—allowing a balancing of results and costs while maintaining DR infrastructure and not losing the momentum that has been built-up so far.

EM&V recognised that an approach and tool that would result in the maximum credibility for its work could equally benefit the program administration of DR itself. Program staff see EM&V and program evaluation partly as a service to them—helping to provide the necessary information to prudently manage their programs. Confidence from system planners is also increased as a result not just of sophisticated results, but because program reporting and projections are so close to *ex post* verified results. One of the goals of most DSM delivery is to have small error bands across the full range of forecast, reported and then EM&V verified energy savings estimates. Clever design of a DR evaluation framework has contributed to this in Ontario for a significant portion of its DSM activities.