

# Reducing Evaluation Uncertainty in Large Custom Projects

*Julianne Meurice, Navigant Consulting, Chicago, IL*

*Nick Beaman, Navigant Consulting, Burlington, VT*

*Lorraine Renta, Navigant Consulting, Verona, WI*

*Jim Jerozal, Nicor Gas, Naperville, IL*

*Scott Dimetrosky, Apex Analytics LLC, Boulder, CO*

*Nathan Warren, CLEAResult, Madison, WI*

## ABSTRACT

Custom commercial and industrial (C&I) projects deliver a significant and growing share of total annual savings in virtually all energy efficiency (EE) portfolios<sup>1</sup>. Because the program savings are not based on pre-determined (deemed) measure gross savings values, the verified project gross savings can deviate significantly – up or down – from the initial program-reported values, and the resulting realization rate (RR) can be significantly different from 1.0. The uncertainty of the net-to-gross ratio (NTGR) further increases the variability of the final program savings. This final savings uncertainty presents a considerable risk for utilities that have state EE savings targets with penalties or rewards attached to their achievement. Frequently, these variations can make the difference between attaining and not attaining the mandated savings goals and paying or not paying the penalty. For utilities without mandated savings goals, there is a lost opportunity to reallocate funds when savings estimates result in either over or under-paying for savings, and program processes that need improvement go unaddressed until after the end of the program year.

This paper presents the story of the collaboration between Nicor Gas and its evaluator to reduce the uncertainty around custom program net savings. Nicor Gas has a three-year net savings target and is subject to \$600,000 in penalties if the target savings are not achieved. Together, we designed two strategic efforts to reduce savings variability, one targeting gross savings' realization rates and one net savings. These initiatives together yielded both less variability in verified savings and faster completion of final reports to the regulator.

## Introduction

Many U.S. EE program providers are challenged to continue to grow their total portfolio savings to meet the increasing stringency of state portfolio savings standards. With the growing maturity of many EE program portfolios, custom commercial and industrial projects are delivering a significant and growing share of total annual portfolio savings as the more common prescriptive measures become more difficult to capture with program maturity.

A key characteristic of these custom projects is that individual project savings are unique to each project and not based on pre-defined measure-based (deemed) gross savings values that are most common in the portfolios. Consequently, the evaluator may affect the program's savings values at two levels, both the gross (utility reported) savings level and the net – that is, after accounting for savings not attributable to the program. Since the evaluator determines final gross project savings values, the

---

<sup>1</sup> This observation reflects Navigant's experience evaluating numerous energy efficiency program portfolios across the U.S. over the past five years.

verified project gross savings can deviate significantly – up or down - from the initial program-reported values and in turn, project realization rates can vary significantly from 1.0. In some states, the evaluator also determines the net-to-gross ratio (NTGR), that is, the ratio of program-attributable savings to verified gross savings. The uncertainty of the NTGR, which for many utilities is unknown until program year-end, only increases the variability of the program savings.

This final program savings estimate uncertainty presents a considerable risk for utilities. Many EE program providers have penalties or rewards tied to the achievement of their state EE savings targets. The custom program savings variations can often make the difference between attaining and not attaining the overall mandated savings goals and paying or not paying the penalty. This reality also leads to considerable discussions about the evaluator’s project specific findings, which may result in delayed final reports and higher evaluation costs. For utilities without such state-mandated savings goals, there is a lost opportunity to reallocate funds when improper savings estimates result in either over or under-paying for savings.

Nicor Gas, an AGL Resources company, operates an energy efficiency portfolio in the state of Illinois that consisted of 14 programs in the three year planning cycle, from 2011 to 2014. Illinois legislation requires that gas utilities achieve 1.2%<sup>2</sup> cumulative savings over the first three-year program cycle. Resource Solutions Group<sup>3</sup> (RSG) implemented the Business Custom Incentive program (Custom Program) and had specific savings targets to enable Nicor Gas’ program portfolio as a whole to meet its savings goals. Navigant Consulting is Nicor Gas’ program evaluator and was responsible for verifying the savings of the Custom Program (and all other EE programs) on both a gross (before adjustment for savings not attributable to the program) and net (after removing savings not attributable to the utility program) basis. Figure 1 presents the typical evaluation process of **custom program gross reported savings**.



**Figure 1.** Typical Custom Program Gross Savings Evaluation Process

The evaluation process follows the international performance measurement and verification protocol (IPMVP®). The IPMVP (EVO 2012), developed by a coalition of international organizations,

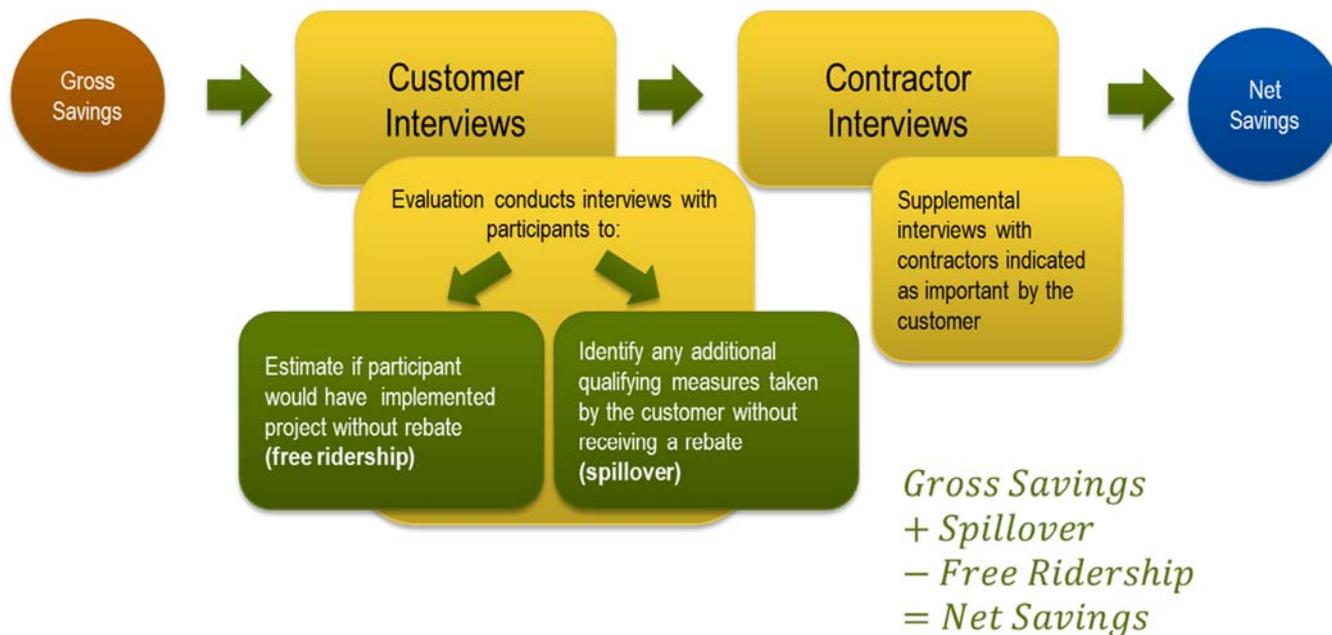
<sup>2</sup> 220 ILCS 5/8-104, Sec. 8-104: (1) 0.2% by May 31, 2012; (2) an additional 0.4% by May 31, 2013, increasing total savings to .6%; (3) an additional 0.6% by May 31, 2014, increasing total savings to 1.2%. The goals continue through 2015-2017. <http://www.ilga.gov/legislation/ilcs/fulltext.asp?DocName=022000050K8-104>

<sup>3</sup> RSG has since been acquired by CLEAResult.

defines standard terms and suggested best practices for quantifying the results of energy efficiency projects (among other objectives). The process involves selecting, after the end of the program year, a sample of completed projects to review the project files and to verify the measures at the customer site. The sample is comprised of three strata classified by project level savings. The results from the projects sampled within each stratum are extrapolated to the remainder of the program's projects that fall within each of the three respective strata. The gross savings verification for this sample involves several activities:

- (1) A review of the project file provided by the implementation contractor (IC) to understand what the IC determined the pre-project (baseline) conditions were, the savings methodology and assumptions for both the baseline and energy efficient equipment, and key project contact information (e.g., customer and contractor).
- (2) For a subset, a visit of the customer site and a visual verification that the specified equipment is installed and operating as intended.
- (3) For this subset, the evaluator also verifies hours of operation, baseline equipment information, and any changes to the facility or equipment process operation (e.g. an increase or decrease in production) through customer interviews at this time.
- (4) Recent billing history is compared to pre-project bills normalized for weather to verify and refine project savings. Alternatively, metering results may be used to verify savings.
- (5) Once the evaluator reports their savings values for each project, frequent communication between the evaluator and IC may occur regarding any varying assumptions. In this scenario, the IC may need to provide additional or updated project-related materials. Repeatedly, the original estimates change to incorporate the new information, and this exchange of information may delay the final program savings report to the regulatory body.

The **typical net savings estimation approach** follows a different process that also begins at the end of the program year. Figure 2 below presents this process.



**Figure 2.** Typical Custom Program Net Savings Estimation Process

The evaluator typically conducts telephone interviews with a sample of participants (usually not the same sample as that used for gross impact research) to understand their decision-making process to

estimate whether they would have done the project without the incentive (free ridership). The interviews also aim to identify any additional measures that they may have taken without obtaining an incentive but that would have qualified (participant spillover). Participant interviews are typically supplemented with interviews with the implementation contractor or vendor when the participant indicates that the contractor was important in their decision-making process. Based upon the interview responses at program year-end, the evaluator develops a final estimate of the percentage of gross verified savings attributable to the program for each project, which is then extrapolated across all program activity.

However, this approach leads to a great deal of uncertainty regarding how custom program verified net savings would vary from reported gross savings. For example, instead of waiting until after the end of the year for net savings values for all projects, the utility would have information at mid-year regarding projects completed in the first quarter and then quarterly thereafter. Additionally, the information gathered earlier in the process is likely to be more accurate as memories of the project decision-drivers will be fresher and the risk the key decision maker has left the company will be reduced. The latter situation also applies to the gross savings estimates in that the lead IC engineer for the project may no longer be available to answer questions regarding key methodological decisions or assumptions, while at the same time the early communication between the evaluator and IC assures that final estimates are more closely aligned than they would have been. Given that Nicor Gas needed to be better able to project its verified net savings to manage effectively the program activities and help avoid penalties for not achieving the mandated savings over the three-year cycle, the evaluator worked with Nicor Gas staff and their IC to modify the method to reduce that uncertainty.

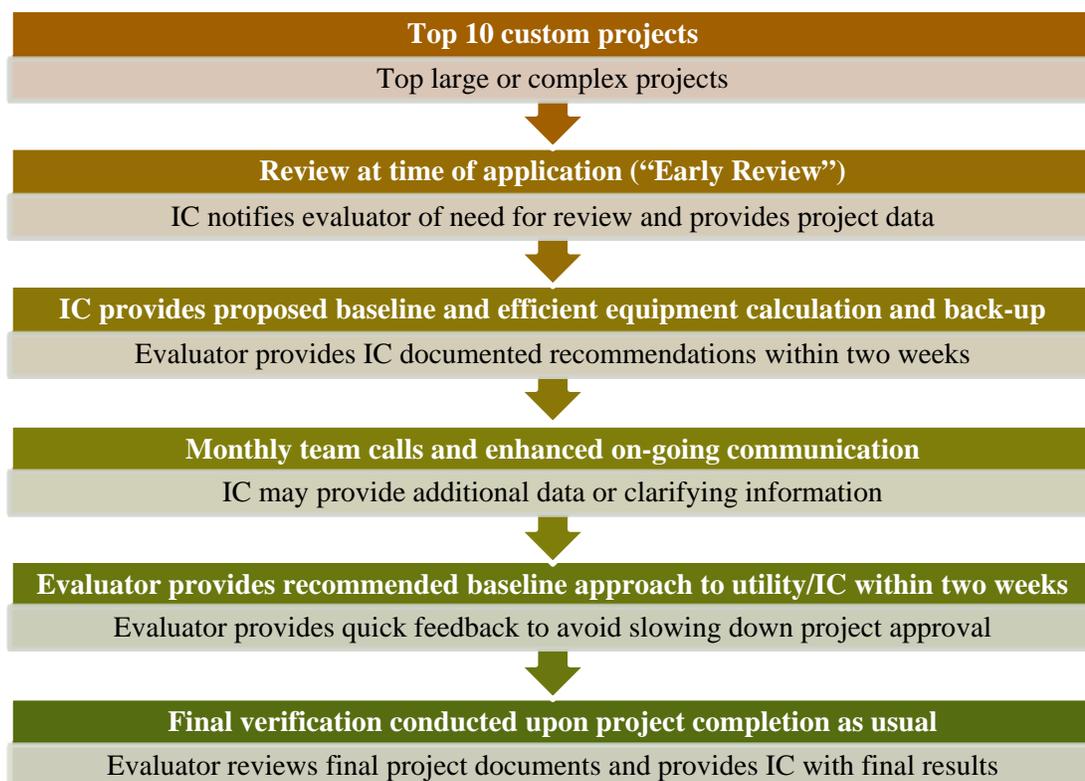
## **Methodology**

The utility, IC and evaluator team (Nicor Gas Team) adopted two processes to reduce the variability in final reported savings, one targeting the gross savings realization rate (RR) and the second the evaluator NTGR compared to the ICs expected NTGR.

### **Reducing Realization Rate Variability**

The Nicor Gas team initially identified the key issue behind the variability in RRs as a baseline issue, specifically the difference in the fundamental assumption as to whether the equipment was being replaced before the end of its life (early replacement) or at the end of its useful life (replaced on burnout). With a closer review of the projects with very high and very low RRs, however, the Nicor Gas team realized that the variance in RRs was driven more by broad issues in savings estimation approaches. Consequently, the team ended up developing a more broad-based approach to address the issues.

Figure 3 details the six key components of the adopted approach used to reduce realization rate variability. The paragraphs below discuss these components.



**Figure 3.** Early Review of Custom Project Gross Savings

### **Top 10 custom projects.**

During the program year, the IC identified 10 projects that had the *potential to generate the largest variance in final gross verified project savings and in total accounted for about 50% of annual program savings*. Specifically, the IC identified two types of projects to focus on: (1) projects with the greatest uncertainty regarding baseline (most notably early replacement or replaced on burn out determination) or (2) the 10 projects anticipated to have most projected first year savings for that program year. The team decided to limit the review to the 10 projects with the highest savings impacts. The team limited the early review scope because of the potential related additional cost as the evaluator will verify these projects again post-implementation.

### **Review at time of application ("Early Review").**

Since the evaluator makes the final determination of project savings, initial savings estimates needed to have more input from the evaluator to reduce differences in the final savings value from the IC's reported savings value. To do this, the IC sent the project files to the evaluator at point of project application (instead of program year-end, after the program cut-off date). This enabled the evaluator to provide input on the appropriate baseline early in the process, before even the estimated incentive was quoted to the potential participant.

Relevant issues addressed at this point in time were not only whether the equipment was appropriately characterized as being replaced early (before it needed to be) or replaced on burn out (because it had to be), but also whether the assumptions and algorithms behind the savings estimation methodology made sense and were correctly applied. In the end, this early review not only enabled the IC to estimate savings more in line with the final evaluated savings but to pay an incentive to the participant that better reflected the verified savings that the utility would realize.

### **IC provides proposed baseline and efficient equipment calculation and back-up.**

At the time of application, the IC provides the evaluator its proposed baseline and efficient equipment calculations as well as the key underlying project files and documentation. This enables the evaluator to develop an early understanding of the project and to begin discussions with the IC regarding their key assumptions and calculations.

### **Monthly team calls and enhanced on-going communication.**

In addition to receiving the project files earlier, the evaluator set up monthly calls for the team to discuss the status of the projects being reviewed and clarify assumptions on both sides. These calls have been very effective in opening up communication on all sides and appear to have had beneficial effects on better aligning the ICs project savings estimates of projects outside the top 10 being reviewed early.

The monthly calls served a multitude of purposes. On these calls the IC identified projects that were ready for review by the evaluator, the evaluator provided a status update on the projects undergoing their review, and allowed the IC and the evaluator to discuss any issues that needed to be addressed. Email and telephone communication between the lead IC (program manager) and IC engineering team and evaluation engineers also enhanced the evaluation process.

Evaluator comments and questions on projects were documented in an early review tracker. This spreadsheet workbook provided detailed comments on the project baseline and savings estimation methodology and documented any concerns with other project documentation, such as incentive levels, application forms, and the like. The implementation contractor then either documented their response to the evaluator's comment or question, or implemented the recommended change to the methodology.

### **Evaluator provides recommended baseline approach to utility/IC within two weeks.**

A key concern about reviewing the project at the time of application is not to slow down project approval for the IC and negatively impact program savings and costs as well as the customer and contractor experience, which is highly valued by the utility. To mitigate this risk, the team agreed that the evaluator needed to provide a recommended baseline approach and savings estimation methodology to the IC no later than two weeks after receiving the full project file, which were documented in an early review project tracker. Upon receiving recommendations from the evaluator, the IC reviews the approach and may set up a meeting with the evaluator to discuss it. If in agreement with the recommendations, the IC incorporates it into a secondary review, including updating baseline estimates and pertinent project conditions.

### **Final verification conducted upon project completion as usual.**

After project completion, the evaluator conducted the final project verification. This final verification can be limited to only a project file review or may also include on-site verification. This will depend on the final sampling approach for all projects completed in the program year. The final verification inspects for "as operated conditions" and may differ from the earlier savings estimate due to a variety of conditions such as changes in operating hours or production.

### **Maintaining the integrity of the evaluation process.**

The authors stress that the integrity of the evaluation process is easily maintained with this process. Both the IC and evaluator understand that the evaluator will fully evaluate the project when it is completed and not assume it is unchanged from its application early review state. The early review process is designed to educate both parties: the evaluator about how the IC views the project's baseline and savings and the IC about how the evaluator will view the baseline and estimate savings. Both parties recognize that there are many other factors that they cannot control that may lead to a realization

rate below or above 1.0; however, these early discussions address fundamental approaches and assumptions where methodological agreement can be expected.

## Improving NTGR Predictability

The second source of variability in net reported savings relates to differences in the ICs expectations regarding customers' decision-making processes and ultimate free ridership levels and what the evaluator learns in their interviews. Typical evaluation timing consists of waiting until the end of the program year to speak to participants. The lag time from when the IC works with the participant to when the evaluator conducts the interviews creates a source of variability. During this time, the individual most knowledgeable about the project may have left the company or position, or have begun to forget many of the project details.

The team's approach to addressing this issue was to conduct two sets of net savings interviews with participants instead of the traditional one interview at program year-end. For projects within a predefined savings threshold, the initial interviews were conducted immediately upon completion of the project to measure free ridership, and the second at program year-end to capture spillover. This free ridership measurement approach has the benefit of increasing the likelihood that participants (as well as any contractors where needed) remember the details of the project and are still with their companies and willing to talk to the evaluator about the project. The follow up interview at year-end allows enough time to have passed for spillover to occur.

## Results

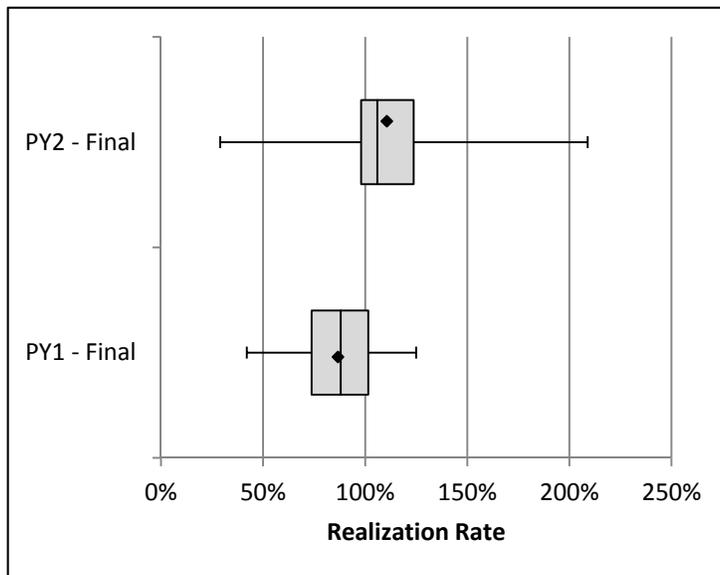
The overall results from the changes in the team's methodologies were quite favorable. The team found four key benefits, which are detailed in this paper:

1. Final RRs moved towards 1.0, enhancing the IC's ability to accurately project program year savings
2. Incentives paid were consequently more in line with the savings attributed to the program
3. Final net savings were more predictable with earlier feedback on selected projects' NTGRs
4. Final program reporting accelerated dramatically due to the on-going communication

Next, the paper compares the processes and related results in program year 1 (PY1), program year 2 (PY2) and program year 3 (PY3) for Nicor Gas' Custom Program evaluation. The evaluation approach for PY1 followed a standard approach, while the evaluation of PY2 and PY3 followed (or is following for PY3), the early review process.

The evaluation results are presented in a box and whisker plot format. The box displays the range of RRs for projects of that year falling in the 25<sup>th</sup> to 75<sup>th</sup> percentiles. The lines at each end of the box display the width to the minimum and maximum RR identified. The line in the middle of the box is the median RR, while the dot is the mean.

**Navigant's final Custom Program RR rose from 0.87 in PY1 (Gunn 2013) to 1.11 in PY2 (Beaman & Renta, 2014), while the range between the 25<sup>th</sup> percentile and 75 percentile RR narrowed** (Figure 4). The range between the minimum and maximum RR were significantly wider in PY2 than in PY1; however, this was due to factors outside the control of the IC, as Figure 6 shows. Specifically, subsequent changes in production schedules and later billing records indicated that savings were either considerably higher or lower than originally estimated.



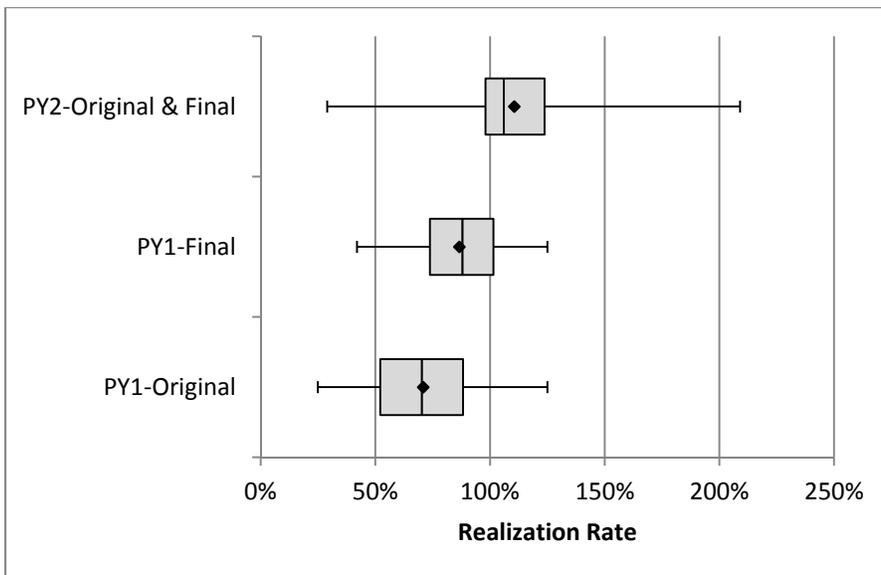
**Figure 4.** Final Realization Rates for PY1 and PY2

The new methodology’s primary objective improved the predictability of final program savings by reducing the variance in final program gross savings. A perfect result would be to achieve a realization rate of 1.0. In PY2, the evaluated gross savings were 11% above the expected level, versus 13% below in PY1. This is not a significant reduction in variance from 1.0; however, the realization rate was also above 1.0. This meant that the utility had achieved more savings towards its goal than anticipated and was in a good position to adjust its plans for the balance of the first three-year portfolio cycle.

These results suggest that the early review process helped the IC better estimate the project findings in line with the evaluator’s final estimates. Of course, there were other factors that could have influenced this outcome including, the IC learning from the PY1 evaluation, a change in IC personnel, or a change in evaluator personnel. Regardless the reason, the outcome was positive for both IC and the utility.

A major reason for a program realization rate greater than 1 in PY2 was the evaluator’s use of a different set of weather data than that used by the IC. In PY3, that variance will be eliminated by using the same data set.

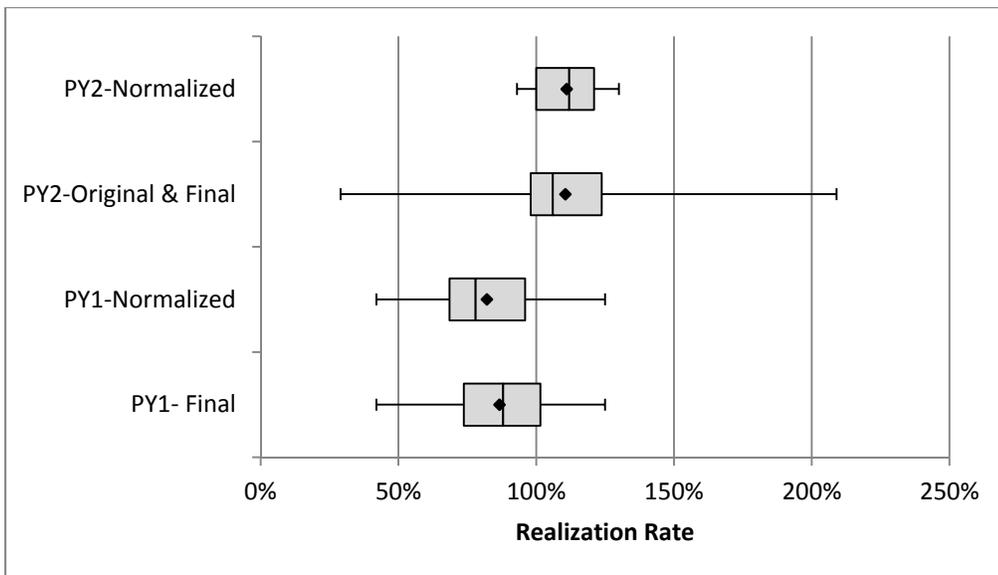
**Navigant’s final Custom Program realization rate in PY1 was 0.87, compared to an initial estimate of 0.71, whereas the final PY2 rate was unchanged at 1.11 (shown in Figure 5).** PY2 results are summarized as “PY2 – Original & Final” while PY1 results are summarized as “PY1 – Final” and “PY1 – Original”. The typical approach (i.e. the PY1 process) to evaluation can require a great deal of communication back and forth between the evaluator and IC to finalize project savings values. For example, if the IC provides incomplete files or data to the evaluator, the IC then resubmits any missing information, necessitating a re-evaluation of the verified savings. With the early review of large custom projects (i.e., the PY2 process), this becomes less necessary for two reasons: there is detailed discussion back and forth on the key projects between the evaluator and the IC before a baseline and savings estimation methodology are agreed upon, and the IC and evaluator are communicating throughout the year on savings and baseline estimation methodologies. This has the additional effect of enhancing the IC’s approaches on most of their other projects.



**Figure 5. PY1 and PY2 Actual Custom Program Realization Rates**

**In parallel, final reporting turnaround time from draft to final report decreased from 10 months in PY1 to less than one month in PY2.** Because the IC and evaluator had been talking with each other over the entire program year, the savings estimation methodologies of the two parties were much more in accord. In addition, the IC had a clear understanding of which documents the evaluator needed to conclude quickly on the savings estimate, eliminating the need for multiple document requests.

**Normalizing the RRs for projects with changes in circumstances out of the IC's control shows more clearly the benefits of the collaboration.** Captured in these realization rates are events that the IC cannot predict that reduce or increase the project savings. Such changes can include significant changes in the operating hours of a customer site or differences in actual versus projected production volumes. To estimate the impact of these uncontrollable events, the evaluator removed from the data sets those projects with clear changes in key parameters outside the IC's control. For example, if a customer's production levels changed after the finalization of the project, and decreased or increased the overall gas usage, the realization rate would be impacted. Because the IC could not have planned around these changes, projects that exhibited uncontrollable events were removed from the analysis. When these projects are removed, the benefits of the team collaboration are even clearer as the normalized PY1 RR is 0.83 compared to a normalized PY2 RR of 1.11. The variability around the median RR is much narrower in PY2 (Figure 6).

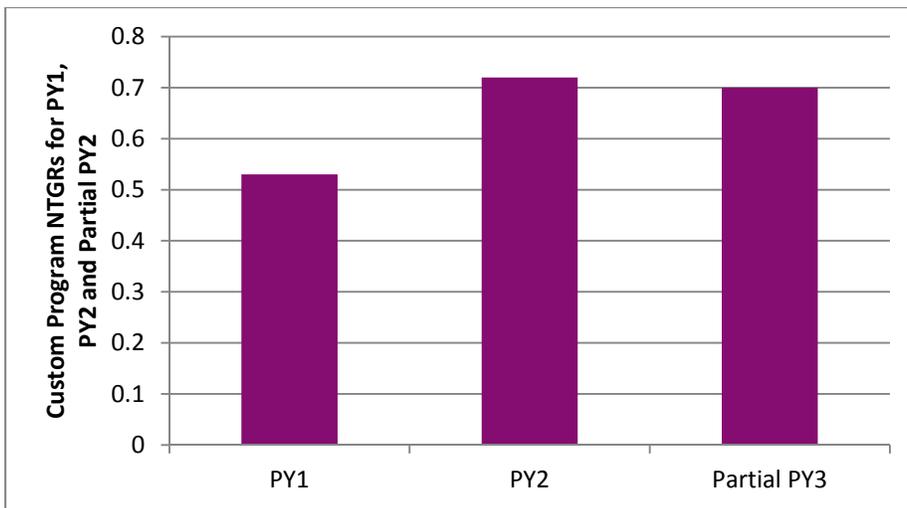


**Figure 6.** Normalized PY2 and PY2 Custom Program Realization Rates

**In PY1 the IC paid for gross savings that the utility did not receive credit for, while the reverse is true in PY2.** A RR of less than 1.0 indicates that the IC overestimated the project savings, and thus could have paid fewer dollars for those savings. The early review process enabled the IC to better estimate project savings and pay for them accordingly. In future years, we would not expect this precise relationship from the early review process (a RR greater than 1.0 instead of less), but rather we would expect the final savings to be more in line with the reported savings, such that incentives paid are more in line with the savings actually generated.

**Early NTGR research in GPY3 has reduced the uncertainty in the final PY3 NTGR and total savings attributed to the utility.** Custom program NTGRs can vary significantly year to year depending on the mix of the technologies and customers and contractors participating. For Nicor Gas' Custom Program, the NTGR ranged from 0.53 in GPY1 to 0.72 in GPY2, a difference of almost 20 percentage points (Figure 7). This difference translates into more or fewer projects needing to be completed to achieve a savings target. Consequently, the evaluator is now conducting NTG research on a sample of custom projects as soon as the project is completed to better enable the IC and utility to conclude on how many projects must be closed by program year-end to meet their net savings target. Going forward, the team plans to interview the customer approximately four weeks after the customer incentive has been approved (which corresponds to the "Pre-Approved" phase of Salesforce<sup>4</sup> Custom pipeline stages) in order to conduct the free-ridership inquiry close to the time of decision-making.

<sup>4</sup> Salesforce is a customer relationship management tool (CRM) utilized by the IC.



**Figure 7.** Custom Program NTGRs for PY1, PY2 and Partial PY3

By the end of April 2014, which is one month before PY3 year-end, Navigant had conducted interviews with seven participants and two of their contractors to obtain an early indication of the PY3 NTGR. These participants' projects represented 67% of total program savings through April 2014 and were projected to account for 18% of total savings for PY3. In contrast, in PY1 and PY2 participants interviewed had completed projects that account for 77% and 81%, respectively of total program savings. Navigant anticipates that we could need to conduct up to another twenty interviews over the next four weeks to complete our PY3 research.

The initial interviews for PY3 yielded a NTGR of 70%, about at the PY2 level and well above the PY1 level. Navigant expects to complete up to another 10 interviews, depending on targeted projects sizes, in the next four weeks to finalize the NTGR, however, so there is still some risk of it changing.

### Potential Issues in the Early Review Process

The Nicor Gas Team's experience is that the potential issues with the early review process are fairly limited and manageable.

- To manage expectations, it is essential that the evaluator state clearly to the utility and the IC at the beginning of the process that the early review does not eliminate all uncertainty and guarantee a RR (or NTGR of 1.0.) The process reduces the verified gross savings surprises, but it does not guarantee specific results since the evaluator will review the project anew under the circumstances in place after project completion. This is some additional upfront cost due to the frequent communication between the IC and evaluator, which has the potential to increase the evaluation cost. An increase is not a certainty, however, because the time-savings at yearend in finalizing the report are considerable and may more than offset the additional communications cost.
- While there is considerable back and forth between the IC and the evaluator about appropriate baselines and savings estimation approaches, it should be clear to the IC that the evaluator has the final say in determining savings levels. In this evaluator's experience, the ICs have understood and accepted this, and it has not been an issue even when there is an extended discussion of the matter. On the contrary, the IC appreciates the opportunity to argue their case.

### Conclusions

The Nicor Gas Team's modifications to the typical impact evaluation approach for custom programs were judged by all the involved parties to be very effective in improving the process for determining the gross savings and net savings that would be attributed to the program. Specifically, this process resulted in nine benefits:

- A higher program realization rate;
- Enhanced ability to manage total project gross savings for PY2;
- More accurate and timely NTGR research results;
- Improved utility satisfaction with the evaluation process;
- Better communications between the evaluator and the implementation contractor;
- Refined technical documentation and clearer documentation requirements
- Timelier, higher quality program feedback;
- Program design changes to better align savings and incentives; and
- More effective use of evaluation resources.

There are additional steps that can and are being taken, however, to reduce that variability still further:

- (1) The team will agree on the appropriate weather data in advance to eliminate that source of difference in savings estimates.
- (2) Navigant will begin to verify the net program impacts on a more real-time basis, that is, begin the impact evaluations earlier in the program year so the verified values – both gross and net – are at least known in part before the end of the program year.

The team is already implementing both additional steps for PY3 and PY4 projects and expects to continue to refine the new approaches throughout the year.

## References

[EVO] Efficiency Valuation Organization, International Performance Measurement & Verification Protocol: Concepts and Options for Determining Energy and Water Savings Volume I, January 2012, [http://www.evo-world.org/index.php?option=com\\_content&view=article&id=272&Itemid=379&lang=en](http://www.evo-world.org/index.php?option=com_content&view=article&id=272&Itemid=379&lang=en)

Gunn, R. 2013. Energy Efficiency Plan Year 1 (6/1/2011- 5/31/2012) Evaluation Report: Nicor Gas Business Custom Incentive Program FINAL, [http://ilsagfiles.org/SAG\\_files/Evaluation\\_Documents/Nicor%20Gas/Nicor%20Gas%20GPY1%20Evaluation%20Reports/Nicor\\_Gas\\_Business\\_Custom\\_Incentive\\_Program\\_GPY1\\_EMV\\_Report%20\\_Final\\_2013.pdf](http://ilsagfiles.org/SAG_files/Evaluation_Documents/Nicor%20Gas/Nicor%20Gas%20GPY1%20Evaluation%20Reports/Nicor_Gas_Business_Custom_Incentive_Program_GPY1_EMV_Report%20_Final_2013.pdf)

Beaman, N. and Renta, L. 2014. Nicor Gas Business Custom Incentive Program GPY2 Evaluation Report Final Energy Efficiency Plan: Gas Plan Year 2 (6/1/2012-5/31/2013), [http://ilsagfiles.org/SAG\\_files/Evaluation\\_Documents/Nicor%20Gas/Nicor%20Gas%20GPY2%20Evaluation%20Reports/Nicor\\_Gas\\_GPY2\\_Custom\\_Eval\\_Report\\_2014-04-02\\_Final.pdf](http://ilsagfiles.org/SAG_files/Evaluation_Documents/Nicor%20Gas/Nicor%20Gas%20GPY2%20Evaluation%20Reports/Nicor_Gas_GPY2_Custom_Eval_Report_2014-04-02_Final.pdf)