

# Peoples Gas and North Shore Gas Home Energy Reports GPY4 – GPY6 Evaluation Plan

## Introduction

This document presents Navigant’s proposed evaluation plan for GPY4<sup>1</sup> of the Peoples Gas (PG) and North Shore Gas (NSG) Home Energy Reports (HER) programs. The HER programs are designed to generate natural gas savings by providing residential customers with information about their specific gas use and related information, as well as conservation suggestions and tips. The information is provided in the form of periodic Home Energy Reports (HERs) that are sent to all participating customers. The HERs give customers three types of information:

- How their recent gas consumption compares to their past gas usage;
- Tips on how they can reduce their gas consumption, some of which are tailored to each customer’s unique circumstances; and
- Information on how their gas use compares to that of neighbors with similar homes.

In GPY3 the HER program was initially rolled out to a targeted sample of 151,200 PG customers and 91,350 NSG customers.<sup>2</sup> In addition, a control group of approximately 21,000 customers were also drawn for each program. In its evaluation of the program in GPY3, Navigant verified average daily savings of 14.2 therms per participant for the PG program, and 7.4 therms per participant for the NSG program. These represented savings rates of 0.85 percent and 0.62 percent of average daily gas usage, respectively, over the October 2013-May 2014 evaluation period. The total verified net program savings in GPY3 was 2,054,727 therms for the PG program and 652,718 therms for the NSG program, compared to GPY savings targets of 900,000 therms and 500,000 therms, respectively.<sup>3</sup>

In GPY4, Opower, the program’s implementation contractor, continued sending HERs to the GPY3 participants who remained part of the recipient sample. No new customers were added to either the recipient groups or the control groups.<sup>4</sup> Thus, in GPY4 most participants will have received reports for at least a full year. For this reason, Navigant expects that average savings per HER recipient will be somewhat greater in GPY4 than in GPY3.<sup>5</sup> On the other hand, holding other factors constant, normal

---

<sup>1</sup> GPY4 began on June 1, 2014 and ended May 31, 2015.

<sup>2</sup> While GPY3 began on June 1, 2013 and ran through May 31, 2014, the first HERs weren’t sent out to participating PG and NSG customers until October 2013. The actual counts of customers who began receiving HERs at that time were 150,152 (PG) and 90,857 (NSG) – personal communication with implementation contractor PM, May 2015.

<sup>3</sup> See Peoples\_Gas\_North\_Shore\_Gas\_GPY3\_Q2\_Report.pdf.

<sup>4</sup> GPY3 HER recipients remained part of the recipient sample in GPY4 unless they opted out of the program or moved. GPY3 control group members remained part of the control sample unless they moved.

<sup>5</sup> Based on previous experience evaluating similar programs, Navigant expects to see a “ramping-up” effect in the first year of a behavioral program, as new participants get used to receiving periodic reports before beginning to take energy-conserving actions in response.

program attrition will tend to lower overall program savings since the average per-customer savings will apply to a smaller base.

## ***Evaluation Objectives***

Navigant proposes to address the following questions in the impact evaluation of the program:

1. What natural gas savings were generated in GPY4 for Peoples Gas and North Shore Gas?
2. How do these savings compare to what was achieved in GPY3?
3. What is the uplift in other Peoples Gas and North Shore Gas energy efficiency programs due to the Opower program?

The GPY4 evaluation will not include a process evaluation since this is a behavioral program by a third-party.

## ***Impact Evaluation Methodology***

The Opower program takes the form of a randomized controlled trial (RCT) in which customers meeting criteria for inclusion in the trial are randomly assigned to either the “treatment” group or the “control” group. Customers in the treatment group receive the HERs, while customers in the control group do not. This experimental design greatly simplifies the estimation of program impacts by providing a group of customers (the controls) whose aggregate energy usage during the program year is an unbiased estimate of the counter-factual case of what the HER recipients would have done in the absence of the program.<sup>6</sup>

The estimation of impacts proceeds in three steps:

1. Validation of the RCT design by comparing the natural gas usage of the treatment and control customers in the pre-program period, to verify that treatment customers are not statistically different than control customers with respect to baseline gas consumption.
2. Econometric analysis of the monthly gas usage of treatment and control customers, described below, to generate estimates of program savings.
3. Statistical analysis to estimate the effect of HER program participation on enrollment uplift in other Peoples Gas and North Shore Gas residential energy efficiency (EE) programs and associated joint savings.

Since Navigant previously validated the randomized design of the GPY3 HER recipient and control samples for its GPY3 evaluation of the program<sup>7</sup>, and no new recipients or controls have been added for GPY4, we will not repeat the validation exercise in GPY4.<sup>8</sup>

---

<sup>6</sup> Randomized assignment from a common pool of eligible customers ensures that the sample characteristics of treatments and controls will be identical except for sampling error. See Paul J. Gertler et al., *Impact Evaluation in Practice*, World Bank 2011, chapter 4.

<sup>7</sup> See PG-NSG Home Energy Reports Program PGY3 Evaluation Final Report, October 3, 2014.

## Econometric analysis to estimate program savings

Navigant will use two regression approaches for estimating program natural gas energy savings for the PG and NSG GPY4 HER program – a simple post-program regression (PPR) analysis with lagged controls and a linear fixed-effects regression (LFER) analysis, applied to monthly billing data. Both approaches produce unbiased estimates of program savings in a RCT, and we will report both sets of results.

However, we expect to use the PPR results for reporting total program savings for GPY4. In our GPY3 evaluation we reported the LFER results, but we have switched this year for two reasons. First, the PPR is the more flexible of the two models, since the method it uses to control for non-treatment differences in usage among customers permits these differences to vary from month to month, unlike the fixed-effects approach. Second, based on our past analyses and those in the academic literature, estimated savings from the PPR model tend to have lower standard errors than those from the LFER model, though the differences are usually small.<sup>9</sup> We will run the LFER model as a robustness check. Although the two models are structurally very different, assuming the RCT is well balanced with respect to the drivers of energy use, in a single sample we expect them to generate very similar estimates of program savings.

### *PPR Model*

The PPR model is shown in equation (1). It controls for non-treatment differences in natural gas use among customers using lagged gas usage as an explanatory variable. Specifically, the model frames the average daily natural gas consumption of a given customer in a given month of the post-program period as a function of whether the customer is receiving HERs or is in the control group, the customer's gas usage in the same month of the pre-program period, a monthly fixed-effect that captures time-varying factors affecting gas usage that are common to all customers (such as weather), and a random error term. The underlying logic is that systematic differences between control and treatment customers will be reflected in differences in their past energy use, which is highly correlated with their current energy use. Formally, the model is the following:

$$(1) \quad ADU_{kt} = \beta_1 Treatment_k + \beta_2 ADU_{lag_{kt}} + \sum_{j=1}^{12} \beta_3_j Month_{jt} + \sum_{j=1}^{12} \beta_4_j Month_{jt} \cdot ADU_{lag_{kt}} + \varepsilon_{kt}$$

where  $ADU_{kt}$  is average daily natural gas use by customer  $k$  in bill period  $t$ ,  $Treatment_k$  is a binary variable which equals one if customer  $k$  is assigned to the group that receives HERs and zero if customer  $k$  is assigned to the control group,  $ADU_{lag_{kt}}$  is customer  $k$ 's gas usage in the same calendar month in the pre-program year as the calendar month of month  $t$ ,  $Month_{jt}$  is a binary variable that equals one when  $j$  equals  $t$  and zero otherwise, and  $\varepsilon_{kt}$  is a random error term.<sup>10</sup>

---

<sup>8</sup> Navigant understands that Opower plans to draw a refill wave in summer of GPY5 for the NSG HER program, which will include approximately 12,000 new HER recipient customers and 10,000 controls, and that the initial HERs will begin going out to the refill recipients in September 2015. Personal communication with implementation contractor PM, May 2015.

<sup>9</sup> Allcott, Hunt and Todd Rogers. "The Short-Run and Long-Run Effects of Behavioral Intervention: Experimental Evidence from Energy Conservation. Forthcoming. *American Economic Review*.

<sup>10</sup> The errors are defined so as to be "cluster-robust," meaning that they are flexible enough to accommodate differing error variances across customers, and correlation over time within customers.

## LFER Model

The LFER model is shown in equation (2). It controls for non-treatment differences in gas consumption among customers using customer-specific fixed effects (intercepts). The LFER model specifies the customer's average daily usage as a function of the customer fixed-effect, a binary variable  $Post_t$  that equals zero if the bill in period  $t$  is received before or in the month of the program start date and one if the bill is received in a month after the program start date, and the interaction between this variable and the binary variable  $Treatment_k$ , which equals one if customer  $k$  is assigned to the group that receives HERs and zero if customer  $k$  is assigned to the control group. Formally,

$$(2) \quad ADU_{kt} = \alpha_{0k} + \alpha_1 Post_t + \alpha_2 Treatment_k \cdot Post_t + \varepsilon_{kt}$$

In this model, the coefficient  $\alpha_{0k}$  captures all customer-specific effects on natural gas usage that do not change over time, such as the square footage and construction of the residence, or the type of HVAC system and hot water heater;  $\alpha_1$  captures the average effect of being in the post-treatment period *across all households in the sample* – for example, weather or economic conditions; while the coefficient  $\alpha_2$  captures the additional effect of being in the post-treatment period specifically for those customers who are in the group receiving HERs – in other words,  $\alpha_2$  measures the mean effect of the program on those customers who are in the program.

Importantly, the experimental design of the program assures that, except for overlap with other Peoples Gas and North Shore Gas energy efficiency programs, this estimate of program savings is an estimate of *net* savings.

## Accounting for Uplift in Other Peoples Gas and North Shore Gas Programs

PPR program savings estimates include savings that result from the uplift in participation in other energy efficiency programs caused by the HER program. To avoid double-counting savings, program savings due to this uplift must be counted towards either the HER program or the other EE programs, but not both. If, during the evaluation period, participation rates in other energy efficiency programs are the same for Opower program participants and controls, the savings estimates from the LFER analysis are already “net” of savings from the other programs, as this indicates the program had no effect on participation in the other energy efficiency programs. If instead the Opower program affects participation rates in other energy efficiency programs, then savings across all programs are lower than indicated by the simple summation of savings in the Opower and energy efficiency programs. For instance, if the Opower program increases participation in one other energy efficiency program, the increase in savings may be allocated to either the Opower program or the energy efficiency program, but cannot be allocated to both programs simultaneously.<sup>11</sup>

A statistic that generates an unbiased estimate of uplift when the baseline average rate of participation in the energy efficiency program is the same for the treatment and control groups –as would be expected (except for random variation) under the randomized controlled trial used to allocate customers across the two groups –is a simple difference in participation rates during the evaluation period. Navigant will use

---

<sup>11</sup> It is not possible to avoid double counting of savings generated by programs for which tracking data is not available, such as upstream CFL programs.

this statistic to generate estimates of participation uplift in other programs, and will use deemed savings for these other programs to estimate joint savings.

### ***Program Evaluation Schedule***

The table below presents an estimate of the evaluation schedule. The schedule for the impact analysis depends on receipt of the necessary data from Opower.

Activity/Deliverables	Responsible Party	Expected Delivery Date
Billing Data to Navigant (June 2014 through May 2015)	Franklin	July 15, 2015
GPY4 Residential Program Tracking Data to Navigant	Franklin	July 15, 2015
Draft Reports Sent to Peoples Gas and North Shore Gas and SAG	Navigant	August 17, 2015
Comments Received on Draft Reports	Franklin, SAG, PG, NSG	August 31, 2015
Final Reports Delivered	Navigant	September 14, 2015

### ***Three Year Evaluation Plan***

At this time, we anticipate the impact evaluation approach for program years GPY5 and GPY6 will repeat the activities of GPY4. We will consider process research if the need arises. Final activities will be determined annually as program circumstances and research questions are better known.