Guidehouse

Residential Education and Outreach Program Impact Evaluation Report

Home Energy Reports Program

Energy Efficiency Plan:
Plan Year 2019
(1/1/2019 - 12/31/2019)

Presented to
Peoples Gas and North Shore Gas

Final

July 8, 2020

Prepared by:

Carly Olig
Guidehouse

Chris Deranian
Guidehouse

www.guidehouse.com

Home Energy Report Impact Evaluation Report

Submitted to:
Peoples Gas
North Shore Gas
200 East Randolph Street
Chicago, IL 60601

Submitted by:

Guidehouse (which acquired Navigant in 2019)
150 N. Riverside Plaza, Suite 2100
Chicago, IL 60606

Contact:

Randy Gunn
Partner
312.583.5714
randy.gunn@guidehouse.com

Kevin Grabner
Associate Director
608.616.5805
kevin.grabner@guidehouse.com

Carly Olig
Associate Director
608.616.4810
carly.olig@guidehouse.com

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Figures

Tables

E. EXECUTIVE SUMMARY

This report summarizes Guidehouse's findings and results from the impact evaluation of Calendar Year 2019 (CY2019)[1] of the Peoples Gas (PGL) and North Shore Gas (NSG) Home Energy Reports (HER) programs. Initially launched in 2013, these programs are designed to generate energy savings by providing residential customers with information about their energy use and energy conservation suggestions and tips. Program participants receive information in the form of paper and email home energy reports and via the customer's energy management portal online.

An important feature of the PGL and NSG HER program is that it is designed as a randomized controlled trial (RCT).[2] Customers in the target group of residential customers from each utility are randomly assigned to either the recipient group or the control (non-recipient) group to estimate changes in energy use due to the program. Customers may opt *out* of the program at any time, but cannot opt *in* due to the RCT design. An implication of the RCT design is that the savings estimates are intrinsically net of free-ridership and most spillover bias. Unless otherwise noted, reported "savings" in this report refer to *net savings*.[3]

In CY2019, the program included three waves of customers: one for PGL and two for NSG. Each utility launched one wave in October 2013 and then restructured their HER programs during 2017 to drop many of the treatment customers from the program. Throughout CY2019, each utility added many of those dropped customers back into the program.[4] NSG also launched a new wave in September 2019. Each treated customer was included in the evaluation only for the period after they (re)started getting reports. Treated customer counts for each of these waves throughout the year are summarized in Table E-1.

Table E-1. CY2019 Peoples Gas and North Shore Gas HER Program Customer Counts

Wave	Treated Customer Counts*		
	2019-01-01	2019-03-01	2019-09-01
PGL	40,000	55,000	81,000
NSG	50,000	59,000	-
NSG – Sept Wave	-	-	18,000

Source: Guidehouse analysis of PGL and NSG customer billing data.
* Customer counts are rounded to the nearest 1000.

[1] CY2019 began January 1, 2019 and ended December 31, 2019.
[2] In selecting each wave, the program implementer, Oracle, randomly allocated targeted PGL and NSG residential customers between participant and control groups. As each wave was added, Guidehouse confirmed that the usage data was consistent with an RCT design.
[3] In some instances, the word "net" appears in column headings and summary sentences for added clarity.
[4] Since these added customers were part of the original RCT design, adding them back into their original program waves does not disrupt the statistical validity of the evaluation.

E.1 Program Savings

Table E-2 summarizes the HER Program's CY2019 natural gas savings. Guidehouse verified net savings of 772,269 therms for PGL and 697,446 therms for NSG after adjustments for persistence from CY2018[5] and uplift[6] resulting in verified net realization rates of 112% and 153%, respectively. Note that savings for the new NSG wave added in September 2019 were not statistically significant resulting in no claimable CY2019 savings for that wave. The persistence adjustment reduced savings by approximately 30% and is also accounted for by the implementer in their *ex ante* savings estimates. The uplift adjustment resulted in approximately a 4% reduction in the net savings which the implementer's savings estimates did not include. The remaining difference in the realization rates were likely due to differences in the regression models used by the evaluation team and the implementer.[7]

Table E-2. CY2019 Peoples Gas and North Shore Gas HER Program Net Savings

Utility	*Ex Ante* Savings (Therms)	Modeled Savings (Therms)*	Verified Savings Prior to Uplift Adjustment (Therms)†	Total Uplift Adjustment (Therms)‡	Verified Net Savings After Uplift Adjustment (Therms)	Verified Realization Rate§
PGL	688,083	1,037,959	825,429	53,160	772,269	112%
NSG	456,028	930,615	713,560	16,114	697,446	153%
NSG – Sept Wave#		15,465	0	-	0	

Source: Guidehouse analysis of PGL and NSG customer billing data.

* Modeled savings are those coming directly out of regression modeling (see Section 2.2) before any adjustments for persistence (see Section 2.3) or uplift (see Section 2.4). The adjustment for persistence is new in CY2019 and reduces the modeled savings by the amount attributable to sending reports in CY2018.[8]

† Verified savings prior to uplift adjust for persistence from CY2018 but not for uplift.

‡ The total uplift adjustment includes both the uplift calculated for CY2019 and the legacy uplift from GPY3 to CY2018. See Section 5.4 for details.

§ The verified realization rate compares verified savings after the uplift adjustment with *ex ante* savings.

Savings for this wave were not statistically significant in CY2019, meaning no savings can be claimed.

[5] The adjustment for persistence is new in CY2019 and reduces the modeled savings by the amount attributable to sending reports in CY2018. This adjustment is prescribed in the *Adjustments to Behavior Savings to Account for Persistence* measure in the Illinois Technical Reference Manual (TRM). See TRM, Measure 6.1.1, Volume 4, Version 7.0. See Section 2.3 for details.

[6] Uplift refers to the impact of the HER program on enrollment in *other* PGL and NSG EE programs. To avoid double-counting the savings from this indirect effect, Guidehouse subtracts the estimated uplift savings from the total HER program savings, including legacy uplift from prior years (see Section 5.3 for details). The fact that uplift savings is subtracted from the HER programs' total energy savings does not indicate that the uplift savings was not *caused by* the HER programs, or that the HER programs shouldn't be credited for its occurrence. It is an accounting adjustment to avoid double-counting when aggregating savings over multiple EE programs. Indeed, the existence of uplift is an indicator of successful cross-marketing by the HER programs, and thus should be seen as an added program benefit. See Section 2.4 for details.

[7] In particular, for the original NSG wave, further review revealed that Guidehouse's and Oracle's modeled estimate of average daily savings per customer were not statistically different from one another.

[8] This adjustment is prescribed in the *Adjustments to Behavior Savings to Account for Persistence* measure in the TRM. See TRM, Measure 6.1.1, Volume 4, Version 7.0.

E.2. Program Volumetric Detail

Table E-3 presents participation details for the CY2019 PGL and NSG HER programs. The PGL wave achieved an average savings rate of 1.00% in CY2019, while the original NSG wave had an average savings rate of 1.19% and the NSG September wave had an average savings rate of 0.22%. Note that savings for the new NSG wave added in September 2019 were not statistically significant resulting in no claimable CY2019 savings for that wave.

Table E-3. CY2019 Peoples Gas and North Shore Gas HER Program Participation Detail

Utility	Number of Participants	Number of Controls	Average Participant Net Savings (Therms)	Average Savings Rate	Average Savings Rate Standard Error
PGL	80,649	13,047	17.15	1.00%	0.20%
NSG	59,203	13,771	16.65	1.19%	0.21%
NSG – Sept Wave*	17,933	14,497	2.62	0.22%	0.22%

Source: Guidehouse analysis of PGL and NSG customer billing data.
* Savings for this wave were not statistically significant in CY2019, meaning no savings can be claimed.

E.3 Findings and Recommendations

For PGL's HER program, Guidehouse verified CY2019 impacts of 772,269 therms and for NSG, the corresponding figure was 697,446 therms.

Finding 1. Average daily savings from the HER program were consistent with previous annual evaluations. As expected, the original waves saw ramp-up from CY2018 into CY2019. Given the expansion of customers added back into the program after a couple years off in CY2019, there could be further ramp-up for these waves in CY2020. As the new NSG wave was only added in September 2019, it is not surprising that savings for that wave were not statistically significant; we expect this wave will ramp-up and have claimable savings in CY2020.

Finding 2. Guidehouse has consistently found greater than 100% realization rates for NSG and, recently, for PGL.

Recommendation 1. If there is interest, Guidehouse could conduct additional analysis comparing Oracle and Guidehouse models to understand the underlying causes of this realization rate difference.

1. INTRODUCTION

This report presents a summary of the findings and results from the impact evaluation of calendar year 2019 (CY2019) for the Peoples Gas (PGL) and North Shore Gas (NSG) Home Energy Reports (HER) program. This program is designed to generate energy savings by providing residential customers with information about their energy use and energy conservation suggestions and tips. Program participants receive information in the form of home energy reports that give customers various types of information, including:

- Assessments of how their recent energy use compares to their own energy use in the past
- Tips on how to reduce energy consumption, some of which are tailored to their own circumstances
- Information on how their energy use compares to that of neighbors with similar homes

Recipient customers received reports by mail and were also invited to log onto a dedicated program website that offers suggestions of additional opportunities to save energy and allows participants to fine-tune their profiles and report conservation steps that they have taken. Other studies have shown that receiving reports containing this type of information can stimulate customers to reduce their energy use, creating average energy savings in the one percent to three percent range, depending on local energy use patterns.

An important feature of the PGL and NSG HER programs is that both were designed as randomized controlled trials (RCTs). Customers in the target group of residential customers from each utility were randomly assigned to either the recipient group or the control (non-recipient) group to estimate changes in energy use due to the program. Having an RCT experimental design makes the process of verifying energy savings simpler and more robust. Among other things, it effectively eliminates free-ridership bias and thus the need for net-to-gross research. Customers may opt *out* of the program at any time, but they cannot opt *in* due to the RCT design.

In CY2019, the program included three waves of customers: one for PGL and two for NSG. Each utility launched one wave in October 2013 and then restructured their HER programs during 2017 to drop many of the treatment customers from the program. Throughout CY2019, each utility added many of those dropped customers back into the program.[9] NSG also launched a new wave (with a new control group) in September 2019. Each treated customer was included in the evaluation only for the period after they (re)started getting reports. Treated customer counts for each of these waves throughout the year are summarized in Table 1-1.

[9] Since these added customers were part of the original RCT design, adding them back into their original program waves does not disrupt the statistical validity of the evaluation.

Table 1-1. CY2019 Peoples Gas and North Shore Gas HER Program Net Savings

Wave	Treated Customer Counts*		
	2019-01-01	2019-03-01	2019-09-01
PGL	42,000	56,000	81,000
NSG	82,000	91,000	-
NSG – Sept Wave	-	-	18,000

Source: Guidehouse analysis of PGL and NSG customer billing data.
* Customer counts are rounded to the nearest 1000.

In its GPY6 evaluation report, Guidehouse confirmed the RCT design of both programs' original waves by comparing the distributions of monthly energy usage of each treatment group-control group pair and verifying that they were consistent with randomized allocation.[10] For the new NSG wave (launched September 2019), Guidehouse confirmed the RCT as part of this year's evaluation (see Section 5.3 for details).

Table 1-2 provides an overview of the number of accounts who received HERs or served as controls along with their average use during the program period. This table shows that the PGL HER recipients used about 20% more natural gas than the NSG participants in the original wave and 30% more than the NSG September wave.

Table 1-2. Synopsis of CY2019 PGL and NSG HER Program Waves

Utility	Number of Participants	Number of Controls	Participant Average Daily Usage in Post Period (Therms)
PGL	80,649	13,047	4.64
NSG	59,203	13,771	3.80
NSG Sept Wave	17,933	14,497	3.31

Source: Guidehouse analysis of PGL and NSG customer billing data.

[10] Navigant, 2019. *Residential Education and Outreach Program Impact Evaluation Report; Home Energy Reports Program. Presented to Peoples Gas and North Shore Gas*

2. EVALUATION APPROACH

The evaluation approach used to produce the results presented in this report is consistent with that of the evaluation in the previous program year, and with evaluations of similar programs in other utilities' territories, relying on statistical analysis appropriate for measuring the impacts of RCTs.

2.1 Data Used in Impact Analysis

In preparation for the impact evaluation, Guidehouse combined and cleaned the data provided by the implementer. Guidehouse performed the following data cleaning steps:

- Filtered data to the pre-period[11] and post period (CY2019) for each wave
- Removed exact duplicate observations
- Aggregated bills that ended in the same month
- Excluded outlier observations, defined as observations with average daily usage outside plus or minus one order of magnitude from the median usage

Detailed accounts of the customers and observations removed by each cleaning step for each wave are included in Section 5.1 of the Appendix.

2.2 Statistical Models Used in the Impact Evaluation

Guidehouse estimated program impacts using two approaches: a lagged dependent variable regression (LDV) analysis with lagged individual controls and a linear fixed-effects regression (LFER) analysis, both applied to monthly billing data. Both approaches should, in principal, produce unbiased estimates of program savings under a wide range of conditions, but Guidehouse prefers the LDV results for two reasons. First, savings estimates produced by the LDV model tend to be more accurate and more precisely estimated than those from the LFER model[12] based on past experience analyzing similar HER programs' impacts and findings from the academic literature.[13] Second, the implementer uses a similar model for their evaluation, which makes the two sets of results comparable. Although the LDV and LFER models are structurally very different, they should generate similar program savings estimates, assuming the RCT is well balanced with respect to the drivers of energy use. Guidehouse used the LDV results for reporting total program savings for CY2019, while the LFER provided a robustness check.

The savings estimates coming out of the LDV regression model are referred to as the modeled savings throughout this report. For final verified savings, the modeled savings must be adjusted for savings persistence (see Section 2.3) and uplift (see Section 2.4).

The LFER model combines cross-sectional and time-series data in a single panel dataset. The regression essentially compares pre- and post-program billing data for participants and controls to identify the effect of the program on usage. The customer-specific fixed effect is a key feature of the

[11] The pre-period differed by wave and in whether the customer was in one of the expansion groups. For the original waves, the pre-period was between October 2012 and September 2013 and for expansion customers encompassed only the months of the pre-period also included in the post period. For the NSG September Wave the pre-period was September to December 2018.

[12] One likely reason for this is that the LDV model embodies more flexibility than the LFER model, in that the former allows the individual customer control variable to vary seasonally while the latter does not – a particularly attractive feature given the highly seasonal nature of natural gas usage. The LFER model treats all unobserved inter-household heterogeneity affecting households' energy usage as time-invariant, while the LDV model uses lagged individual controls that can vary over time. This is discussed in more detail in Section 5.2 of the Appendix.

[13] Allcott, Hunt and Todd Rogers, 2014. "The Short-Run and Long-Run Effects of Behavioral Intervention: Experimental Evidence from Energy Conservation." *American Economic Review*, 104(10): 3003-37.

LFER analysis and captures all customer-specific factors affecting natural gas usage that do not change over time, including those that are unobservable. Examples of the latter include the construction and square footage of the premise, the number of occupants, the amount of seasonal sun exposure, and the thermostat settings. The fixed effect represents an attempt to control for any small, systematic differences between the treatment and control customers that might occur due to chance.

Like the LFER model, the LDV model also combines cross-sectional and time-series data in a panel dataset. Unlike the LFER model, however, it uses only the post-program data in the dependent variable and includes the customer's lagged energy usage for the same calendar month of the pre-program period to serve as the control for any small, systematic differences between the treatment and control customers, in that sense serving the same purpose as the customer fixed effect included in the LFER model. Section 5.2 of the Appendix presents the details of the LDV and LFER models used in the analysis.

2.3 Accounting for Savings Persistence

Continued implementation of HER programs in Illinois and across the country has demonstrated persistence of savings beyond the first year leading Illinois to adopt a measure persistence framework in Version 7.0 of the TRM. This framework assumes that savings persist over five years but the persistence decays in each year. The TRM recommends using the persistence factors presented in Table 2-1 over the five-year life to estimate lifetime gas savings for the program. In CY2019, the original PGL and NSG waves are in Year 2 while the new NSG September wave is in Year 1.

Table 2-1. Residential Behavior Electric Savings Persistence Factors

Year	Gas Persistence Factor
Year 1	100%
Year 2	45%
Year 3	20%
Year 4	9%
Year 5	4%

Source: TRM, Measure 6.1.1, Volume 4, Version 7.0

Per the TRM, the adjustment for persistence also accounts for the program retention rate. The retention rate was based on the treatment customers who were also in the program in CY2018 (i.e., not the treatment customers added into the program in the March and September expansions). The retention rates were 89.6% and 82.8% for NSG and PGL, respectively.

2.4 Accounting for Uplift in Other Energy Efficiency Programs

The home energy reports sent to participating households included energy-saving tips, some of which encouraged participants to enroll in other PGL-NSG EE programs. If participation rates in other EE programs were the same for HER participant and control groups, the savings estimates from the regression analysis are already "net" of savings from the other programs, as this indicates the HER Program had no net effect on participation in the other EE programs. However, if the receipt of reports increased participation rates of recipients relative to controls in other EE programs, then the combined savings across all programs would be lower than indicated by the simple summation of savings in the

HER and the other EE programs. For instance, if the HER Program increases participation in another EE program, the resulting increase ("uplift") in savings may be allocated to either the HER Program or the EE program, but cannot be allocated to both programs simultaneously.[14]

As data permitted, Guidehouse used a difference-in-difference (DID) statistic to estimate uplift in other EE programs. To calculate the DID statistic, Guidehouse calculated the difference between the HER treatment and control groups in average EE program savings per customer in the post period,[15] and subtracted the same difference from the pre-period. For instance, if the EE program savings during CY2019 is five therms for the treatment group and three therms for the control group, and the savings during the year before the start of the HER Program is two therms for the treatment group and one therm for the control group, then the DID statistic is one therm, as reflected the following calculation:

(CY2019 treatment group savings – CY2019 control group savings) – (pre-year treatment group savings - pre-year control group savings) = DID statistic

$$(5 - 3) - (2 - 1) = 1$$

The DID statistic generates an unbiased estimate of uplift when the baseline average savings is the same for the treatment and control groups, or when they are different due only to differences between the two groups in time-invariant factors, such as the square footage of the residence.

An alternative statistic that generates an unbiased estimate of uplift when the baseline average savings in the EE program is the same for the treatment and control groups is a simple difference in savings during CY2019. Guidehouse uses this alternative statistic –the "post-only difference" (POD) statistic – in cases where the EE program did not exist for the entire pre-program year.

Note that this approach differs slightly from the approach Guidehouse has used to account for uplift in years past. This approach relies on the difference in *savings* from other EE programs between the HER treatment and control groups, rather than the difference in *participation*. Since most programs encompass multiple measures or intensity of measures, this approach better captures any differences in *how* HER participants participate in other programs rather than just *whether* they participate (for example, HER participants could save more from weatherization than HER controls because they install more insulation).

Guidehouse examined the uplift associated with seven other PGL-NSG EE programs: Home Energy Jumpstart (HEJ), Home Energy Rebate (HEReb), Multifamily Energy Savings (MF),[16] Income-Eligible Single-Family (SFIE),[17] Income-Eligible Weatherization (WXIE), and Income-Eligible Multi-Family

[14] It is not possible to avoid double-counting of the savings generated by programs for which tracking data are not available, such as upstream lighting programs.

[15] Where the averages are calculated over all treatment and control group customers, not just those who participated in other EE programs.

[16] MF includes direct install, custom, partner trade ally, and prescriptive.

[17] SFIE includes Chicago Bungalow Association (CBA) and Illinois Home Weatherization Assistance Program (IHWAP).

(MFIE).[18,19] For each EE program, uplift savings were calculated separately for each utility. In addition, legacy uplift (uplift from CY2018, GPY6, GPY5, GPY4, and GPY3) was also calculated.[20]

2.5 Process Evaluation

Guidehouse's CY2019 PGL and NSG HER process evaluation included interviews with the program implementer to update our information about the program, such as plans for additional waves. The evaluation did not include any participant surveys or interviews.

[18] MFIE includes IHWAP and Income-Eligible Multifamily Savings (IEMS).

[19] Guidehouse also looked at the Public Housing Energy Savings program, but found no overlap with HER. Additionally, double counting between the Affordable Housing New Construction Programs and HER is not possible due to the requirement that HER participants have sufficient historical usage data.

[20] Legacy uplift refers to uplift from prior years for which the measure life of the applicable program has not yet passed. These savings are also de-rated by the average move out rate to account for savings which no longer get captured in our HER analysis. Guidehouse used a move out rate of 6%.

3. NET IMPACT EVALUATION

A key feature of the RCT design of the HER program is that the analysis inherently estimates net savings because there are no participants who would have received the individualized reports in the absence of the program. While some customers receiving reports may have taken energy-conserving actions or purchased high-efficiency equipment anyway, the random selection of program participants (as opposed to voluntary participation) implies that the control group of customers not receiving reports would be expected to exhibit the same degree of energy-conserving behavior and purchases. Therefore, this method estimates net savings, and no further NTG adjustment is necessary.

Table 3-1 summarizes total program savings. Note that savings for the new NSG wave added in September 2019 were not statistically significant resulting in no claimable CY2019 savings for that wave. Relative to the implementer's *ex ante* estimates, Guidehouse verified higher savings for both PGL and NSG.[21] Guidehouse adjusted the modeled savings estimated from the regression models by persistence from CY2018[22] and uplift.[23]

[21] Further investigation suggests that the differences in results between Oracle and Guidehouse were mainly driven by small differences in the regression modeling. In particular, for the original NSG wave, further review revealed that Guidehouse's and Oracle's modeled estimate of average daily savings per customer were not statistically different from one another.

[22] The adjustment for persistence is new in CY2019 and reduces the modeled savings by the amount attributable to sending reports in CY2018. This adjustment is prescribed in the *Adjustments to Behavior Savings to Account for Persistence* measure in the TRM. See TRM, Measure 6.1.1, Volume 4, Version 7.0. See Section 2.3 for details.

[23] Uplift refers to the impact of the HER program on enrollment in *other* PGL and NSG EE programs. To avoid double-counting the savings from this indirect effect, Guidehouse subtracts the estimated uplift savings from the total HER program savings, including legacy uplift from prior years (see Section 5.3 for details). The fact that uplift savings is subtracted from the HER programs' total energy savings does not indicate that the uplift savings was not *caused by* the HER programs, or that the HER programs shouldn't be credited for its occurrence. It is an accounting adjustment to avoid double-counting when aggregating savings over multiple EE programs. Indeed, the existence of uplift is an indicator of successful cross-marketing by the HER programs, and thus should be seen as an added program benefit. See Section 2.4 for details.

Table 3-1. CY2019 PGL and NSG HER Program Gas Savings

Utility	*Ex Ante* Savings (Therms)	Modeled Savings (Therms)*	Verified Savings Prior to Uplift Adjustment (Therms)†	Total Uplift Adjustment (Therms)‡	Verified Net Savings After Uplift Adjustment (Therms)	Verified Realization Rate§
PGL	688,083	1,037,959	825,429	53,160	772,269	112%
NSG	456,028	930,615	713,560	16,114	697,446	153%
NSG – Sept Wave#		15,465	0	-	0	

Source: Guidehouse analysis of PGL and NSG customer billing data.

* Modeled savings are those coming directly out of regression modeling (see Section 2.2) before any adjustments for persistence (see Section 2.3) or uplift (see Section 2.4). The adjustment for persistence is new in CY2019 and reduces the modeled savings by the amount attributable to sending reports in CY2018.[24]

† Verified savings prior to uplift adjust for persistence from CY2018 but not for uplift.

‡ The total uplift adjustment includes both the uplift calculated for CY2019 and the legacy uplift from GPY3 to CY2018. See Section 5.4 for details.

§ The verified realization rate compares verified savings after the uplift adjustment with *ex ante* savings.

Savings for this wave were not statistically significant in CY2019, meaning no savings can be claimed.

3.1 LDV and LFER Model Parameter Estimates

The LDV and LFER models generated very similar results for program savings estimates for the two NSG waves. For the PGL wave, the LDV and LFER model were fairly different; the LDV model gave savings within the expected range while the LFER model predicted statistically significant dis-savings (i.e., an increase in usage for the treatment customers). As outlined in Section 2.2 Guidehouse prefers the LDV model and used the LDV results for reporting CY2019 total program savings for all waves.[25]

3.2 Uplift Analysis Results

The LDV estimates include savings that resulted from participation in other EE programs caused by the HER program. To avoid double-counting when aggregating savings across the portfolio, Guidehouse removes from HER impacts uplift in other EE programs. Legacy uplift captures energy savings from previous program years (GPY3, GPY4, GPY5, GPY6, and CY2018) for measures that have multi-year measure lives. CY2019 uplift captures savings from other EE programs that occurred in 2019. Table 3-2 shows uplift figures for PGL and NSG, and how the adjustment affected total savings. Note that savings for the new NSG wave added in September 2019 were not statistically significant resulting in no claimable CY2019 savings for that wave; numbers shown in this table show Guidehouse's estimated savings for reference even though they are not claimable.

[24] This adjustment is prescribed in the *Adjustments to Behavior Savings to Account for Persistence* measure in the TRM. See TRM, Measure 6.1.1, Volume 4, Version 7.0.

[25] Because of the different results between the two models, Guidehouse also tested a post only model for the PGL wave, which estimated savings very close to the LDV model. The post only model only used 2019 data and regressed usage on a treatment indicator and a set of monthly dummies.

Table 3-2. CY2019 PGL and NSG Uplift Results

	PGL Savings (Therms)	NSG Savings (Therms)	NSG Sept Wave Savings (Therms)
Net Savings, Prior to Uplift Adjustment	825,429	713,560	15,465
CY2019 Uplift Adjustment	38,269	3,265	751
Legacy Uplift Adjustment	14,891	12,849	0
Final Net Savings	772,269	697,446	14,714

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

Section 5.4 in the Appendix presents detailed calculations of CY2019 and legacy uplift for each of the EE programs considered in the analysis: HEJ, HEReb, MF, SFIE, WXIE, and MFIE.

3.3 Verified Program Impact Results

Table 3-3 summarizes estimated program savings by participant wave, including CY2019 and legacy uplift adjustments. The table also includes the number of participants, controls, and average savings rates. Both modeled savings and average savings rates include standard error figures. Note that savings for the new NSG wave added in September 2019 were not statistically significant with 90% confidence resulting in no claimable CY2019 savings for that wave; numbers shown in this table show Guidehouse's estimated savings for reference even though they are not claimable.

Table 3-3. PGL and NSG CY2019 HER Program Savings

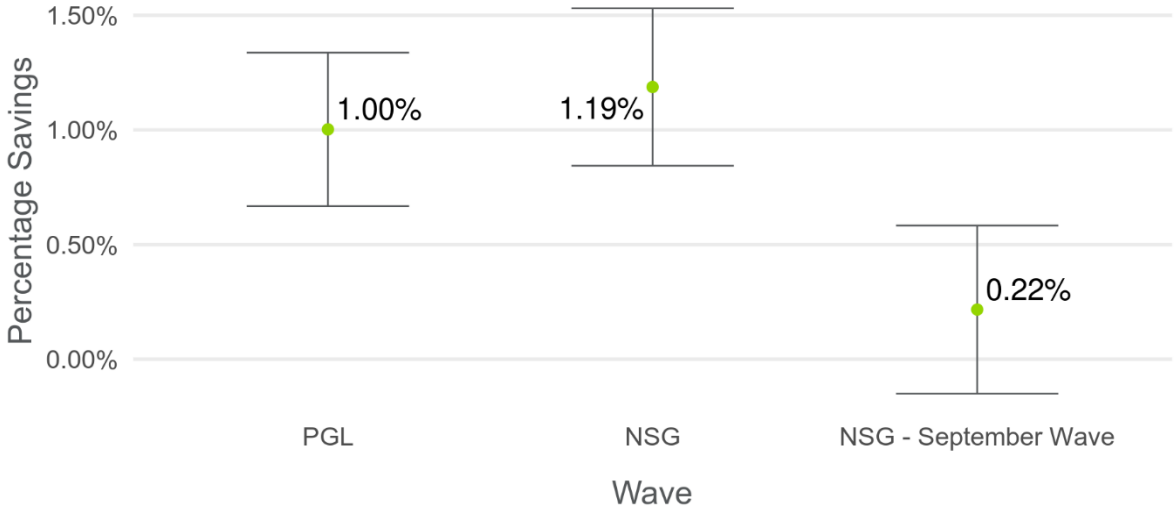
Savings Category	PGL	NSG	NSG – Sept Wave*
Ex Ante Net Savings, therms	688,083		456,028
Number of Participants	80,649	59,203	17,933
Number of Controls	13,047	13,771	14,497
Modeled Savings, therms	1,037,959	930,615	15,465
(Standard Error)	210,658	163,600	15,924
Average Savings Rate (%)	1.00%	1.19%	0.22%
(Standard Error)	0.20%	0.21%	0.22%
Net Savings Prior to Uplift Adjustment, therms	825,429	713,560	15,465
CY2019 Uplift Adjustment, therms	38,269	3,265	751
Legacy Uplift, therms	14,891	12,849	0
Total Uplift Adjustment, therms	53,160	16,114	751
Net Savings After Uplift Adjustment, therms	772,269	697,446	14,714

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

* Savings for this wave were not statistically significant in CY2019, meaning no savings can be claimed.

Figure 3-1 shows energy savings for each wave with 90% confidence intervals.

Figure 3-1. CY2019 Percent Savings and 90% Confidence Interval, by Wave



Source: Guidehouse analysis of PGL and NSG customer billing data.

4. FINDINGS AND RECOMMENDATIONS

For PGL's HER program, Guidehouse verified CY2019 impacts of 772,269 therms and for NSG, the corresponding figure was 697,446 therms.

Finding 1. Average daily savings from the HER program were consistent with previous annual evaluations. As expected, the original waves saw ramp-up from CY2018 into CY2019. Given the expansion of customers added back into the program after a couple years off in CY2019, there could be further ramp-up for these waves in CY2020. As the new NSG wave was only added in September 2019, it is not surprising that savings for that wave were not statistically significant; we expect this wave will ramp-up and have claimable savings in CY2020.

Finding 2. Guidehouse has consistently found greater than 100% realization rates for NSG and, recently, for PGL.

Recommendation 1. If there is interest, Guidehouse could conduct additional analysis comparing Oracle and Guidehouse models to understand underlying causes of this realization rate difference.

Historical Results

Table 4-1 below shows the historical net savings realization rates for the HER Program. The impact analysis method provides net savings directly. Gross savings are not estimated, and there is no NTG ratio.

Table 4-1. Historical Realization Rates and NTG Values

Program Year	PGL Verified Net Savings RR	NSG Verified Net Savings RR	PGL NTG	NSG NTG
GPY1	No Program	No Program		
GPY2	No Program	No Program		
GPY3	105%	98%	NA	NA
GPY4	110%	125%	NA	NA
GPY5	98%	101%	NA	NA
GPY6	92%	116%	NA	NA
2018	106%	129%	NA	NA
2019	112%	153%	NA	NA

Source: Guidehouse evaluation research. Analysis method provides net savings directly. The program was first offered in GPY3.

5. APPENDIX 1. IMPACT METHODOLOGY DETAIL

5.1 Detailed Data Cleaning

Guidehouse performed the following data cleaning steps:

- Excluded post-period data from outside of the period of examination (calendar year 2019)
- Filtered to relevant pre-period data for each wave
- Removed exact duplicate observations
- Aggregated bills that ended in the same month
- Excluded outlier observations, defined as observations with average daily usage outside plus or minus one order of magnitude from the median
- For the LDV model, removed observations that did not have a usage value in the same month of the pre-period.

Table 5-1, Table 5-2, and Table 5-3 give counts of customers and observations removed for the data cleaning steps identified above. Each data cleaning step removed a similar percentage of treatment and control customers for each wave. This suggests that non-random biases were not introduced into the data by the cleaning steps.

Table 5-1. NSG CY2019 Data Cleaning Results

Cleaning Step	Customers		Observations	
	Treatment	Control	Treatment	Control
Raw Data	59,203	13,771	5,233,409	1,225,906
Subset to pre/post periods	59,203	13,771	1,332,506	318,972
Remove exact duplicate observations	59,203	13,771	1,332,506	318,972
Bill Flattening	59,203	13,771	1,296,977	310,172
Exclude outliers	59,203	13,771	1,294,983	309,598
Remove pre-period data (for LDV analysis)	58,070	13,520	630,977	150,954
Remove observations without a monthly pre-use value (for LDV analysis)	58,051	13,514	610,077	145,973

Source: Guidehouse analysis of PGL and NSG customer billing data.

Table 5-2. NSG September Wave CY2019 Data Cleaning Results

Cleaning Step	Customers		Observations	
	Treatment	Control	Treatment	Control
Raw Data	17,933	14,497	336,571	271,958
Subset to pre/post periods	17,928	14,489	278,320	224,820
Remove exact duplicate observations	17,928	14,489	278,320	224,820
Bill Flattening	17,928	14,489	269,986	218,070
Exclude outliers	17,928	14,489	269,371	217,608
Remove pre-period data (for LDV analysis)	17,790	14,379	64,391	51,915
Remove observations without a monthly pre-use value (for LDV analysis)	17,726	14,308	59,994	48,355

Source: Guidehouse analysis of PGL and NSG customer billing data.

Table 5-3. PGL CY2019 Data Cleaning Results

Cleaning Step	Customers		Observations	
	Treatment	Control	Treatment	Control
Raw Data	80,649	13,047	6,658,723	1,151,898
Subset to pre/post periods	80,647	13,047	1,391,005	299,785
Remove exact duplicate observations	80,647	13,047	1,391,005	299,785
Bill Flattening	80,647	13,047	1,363,363	293,625
Exclude outliers	80,647	13,047	1,363,290	293,608
Remove pre-period data (for LDV analysis)	76,491	12,384	654,498	141,615
Remove observations without a monthly pre-use value (for LDV analysis)	76,455	12,384	635,341	137,693

Source: Guidehouse analysis of PGL and NSG customer billing data.

5.2 Detailed Impact Methodology

Guidehouse used two regression models to estimate impacts: an LDV model and an LFER model. The following sections present each model.

5.2.1 LDV Model

The LDV model controls for non-program differences in energy use between the treatment and control groups using each customer's lagged energy usage as an explanatory variable. In particular, the model frames energy use in calendar month *t* of the post-program period as a function of both the treatment variable and energy use in the same calendar month of the pre-program period. 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 shown in Equation 5-1.

Equation 5-1. Lagged Dependent Variable Regression Model

$$ADU_{kt} = \beta_1 Treatment_k + \sum_J \beta_{2j} Month_{jt} + \sum_J \beta_{3j} Month_{jt} \cdot ADUlag_{kt} + \varepsilon_{kt}$$

where:

ADU_{kt}	is average daily consumption of therms by household k in bill period t
$Treatment_k$	is a binary variable taking a value of 0 if household k is assigned to the control group, and 1 if assigned to the treatment group
$ADUlag_{kt}$	is household k's energy use in the same calendar month of the pre-program year as the calendar month of month t
$Month_{jt}$	is a binary variable taking a value of 1 when $j = t$ and 0 otherwise[26]
ε_{kt}	is the cluster-robust error term for household k during billing cycle t; cluster-robust errors account for heteroscedasticity and autocorrelation at the household level.[27]

The coefficient β_1 is the estimate of the average daily therms energy savings due to the program.

5.2.2 LFER Model

The LFER model used by Guidehouse is one in which average daily consumption of therms by household k in bill period t, denoted by ADU_{kt}, is a function of the following three terms:

1. The binary variable $Treatment_k$
2. The binary variable $Post_t$, taking a value of 0 if month t is in the pre-treatment period, and 1 if in the post-treatment period.
3. The interaction between these variables, $Treatment_k \cdot Post_t$

Formally, the LFER model is shown in Equation 5-2.

Equation 5-2. Linear Fixed Effects Regression Model

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

In this model, the coefficient α_{0k} captures all household-specific effects on energy use that do not change over time, including those that are unobservable, the coefficient α_2 captures the average effect *across all households* of being in the post-treatment period, and the effect of being both in the treatment group and in the post period (i.e., the effect directly attributable to the program) is captured by the coefficient α_2. In other words, while the coefficient α_1 captures the change in average daily therms use across the pre- and post-treatment for the *control* group, the sum α_1 + α_2 captures this change for the treatment group, and so α_2 is the estimate of average daily therms energy savings due to the program.

[26] In other words, if there are *T* post-program months, there are *T* monthly dummy variables in the model, with the dummy variable $Month_{tt}$ the only one to take a value of 1 at time t. These are, in other words, monthly fixed effects.

[27] Ordinary Least Squares (OLS) regression models assume that the data are homoscedastic and not autocorrelated. If either of these assumptions is violated, the resulting standard errors of the parameter estimates are incorrect (usually underestimated). A random variable is heteroskedastic when the variance is not constant. A random variable is autocorrelated when the error term in one period is correlated with the error terms in at least some of the previous periods.

5.3 New Wave RCT Checks

To test that the new NSG September 2019 wave is consistent with an RCT, Guidehouse compared treatment and control usage for each month during the pre-program period (September 2018 to August 2019). If the allocation of households across participants and controls is truly random, the two groups should have the same distribution of usage during the twelve months prior to receiving the program intervention. The evaluation team conducted variance tests and t-tests comparing participant and control usage for each month of the pre-period and found that mean usage was not statistically different. As an additional check, the evaluation team performed a regression analysis in which average daily usage in the pre-program period was a function of monthly binary variables and a binary participation variable which showed participation did not impact usage.

Figure 5-1 illustrates control group and report recipient group usage during the twelve-month pre-period. This graph illustrates what Guidehouse's statistical analysis confirmed, namely that the assignment of customers into the treatment and control groups was consistent with randomization.

Figure 5-1. RCT Usage Comparison for NSG September Wave

Source: Guidehouse analysis of NSG customer billing data.

5.4 Detailed Uplift Analysis Results

5.4.1 CY2019 Uplift

Table 5-4, Table 5-5, and Table 5-6 present program savings due to participation in other EE programs in CY2019. Each table provides the uplift for a single program group in each of six EE Programs for which estimates for deemed savings are available.[28] If a particular EE program does not show up in the table for a given wave, it means that HER wave had no participation in that EE program. While these tables show estimates of both positive and negative uplift, only positive values were used to adjust program savings for double-counting. For all cases where the EE program did not exist in the pre-program year, the estimate is based on a POD statistic; otherwise it is based on a DID statistic.[29]

Table 5-4. CY2019 PGL HER Uplift Adjustment Details

Program	HEJ	HEReb	MF	MFIE	SFIE	WXIE
Median program savings (annual therm per EE participant)	49.85	276.00	28.59	527.79	380.59	36.17
Number of treatment customers	80,664	80,664	80,664	80,664	80,664	80,664
Number of control customer	13,050	13,050	13,050	13,050	13,050	13,050
Avg savings per HER treatment customer, CY19	0.14	0.09	0.00	0.01	0.36	0.01
Avg savings per HER control customer, CY19	0.22	0.15	0.00	0.00	0.47	0.02
CY19 savings difference	-0.08	-0.05	0.00	0.01	-0.12	0.00
Avg savings per HER treatment customer, pre	0.45	2.24	0.00	0.00	0.00	0.00
Avg savings per HER control customer, pre	0.52	2.76	0.01	0.00	0.00	0.00
Pre savings difference	-0.06	-0.52	0.00	0.00	0.00	0.00
DID or POD statistic	-0.02	0.46	0.00	0.01	-0.12	0.00
Savings attributable to other programs (therm)	-1287	37,340	216	713	-9,359	-339
Implied change in participation	-25.8	135.3	7.6	1.4	-24.6	-9.4

Source: Guidehouse analysis of PGL program tracking and customer billing data.

[28] See Section 2.4 for more information about the programs considered.

[29] See Section 2.4 for more information on POD and DID statistics.

Table 5-5. CY2019 NSG HER Uplift Adjustment Details

Program	HEJ	HEReb	MF	MFIE	SFIE
Median program savings (annual therm per EE participant)	36.95	272.00	31.31	163.25	262.39
Number of treatment customers	91,350	91,350	91,350	91,350	91,350
Number of control customer	21,000	21,000	21,000	21,000	21,000
Avg savings per HER treatment customer, CY19	0.10	0.35	0.00	0.04	0.00
Avg savings per HER control customer, CY19	0.09	0.36	0.00	0.04	0.00
CY19 savings difference	0.01	-0.01	0.00	0.00	0.00
Avg savings per HER treatment customer, pre	0.08	2.50	0.02	0.00	0.00
Avg savings per HER control customer, pre	0.09	2.35	0.02	0.00	0.00
Pre savings difference	-0.02	0.15	-0.01	0.00	0.00
DID or POD statistic	0.03	-0.16	0.01	0.00	0.00
Savings attributable to other programs (therm)	2,728	-14,205	490	19	29
Implied change in participation	73.8	-52.2	15.6	0.1	0.1

Source: Guidehouse analysis of NSG program tracking and customer billing data.

Table 5-6. CY2019 NSG September Wave HER Uplift Adjustment Details

Program	HEJ	HEReb	MFIE	SFIE
Median program savings (annual therm per EE participant)	22.93	131.50	175.89	174.09
Number of treatment customers	17,975	17,975	7,975	17,975
Number of control customer	14,535	14,535	14,535	14,535
Avg savings per HER treatment customer, CY19	0.01	0.07	0.00	0.00
Avg savings per HER control customer, CY19	0.02	0.08	0.00	0.00
CY19 savings difference	-0.01	-0.01	0.00	0.00
Avg savings per HER treatment customer, pre	0.08	0.54	0.03	0.03
Avg savings per HER control customer, pre	0.06	0.50	0.05	0.05
Pre savings difference	0.01	0.04	-0.02	-0.02
DID or POD statistic	-0.03	-0.04	0.02	0.02
Savings attributable to other programs (therm)	-493	-778	356	395
Implied change in participation	-21.5	-5.9	2.0	2.3

Source: Guidehouse analysis of NSG program tracking and customer billing data.

5.4.2 Legacy Uplift

To determine legacy uplift, Guidehouse utilized uplift in other PGL-NSG EE programs from previous evaluation. The total resource cost report provided individual measure lives.[30] They are the simple average of the measures included in that program. Table 5-7 to Table 5-11 show double counted savings

[30] Navigant Consulting, 2016. *Plan Year 1 through 3 Total Resource Cost Test Results and Impact Summary Evaluation Report.* Presented to Peoples Gas

(therms) from each program for GPY3 to CY2018. These tables show estimates of both positive and negative uplift; however, only positive uplift was used to adjust program savings for double-counting. Additionally, the adjustment for these savings included a de-rating of 6% to account for attrition from the HER program.

Table 5-7. Doubled Counted Savings (Therms) from GPY3

	HEJ	HEReb	MF
Measure Life	10	15	12
PGL	78	-10,222	3
NSG	2,503	4,195	292

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

Table 5-8. Doubled Counted Savings (Therms) from GPY4

	HEJ	HEReb	MF
Measure Life	10	15	12
PGL	1,261	-5,067	227
NSG	1,085	-50,262	321

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

Table 5-9. Doubled Counted Savings (Therms) from GPY5

	HEJ	HEReb	MF
Measure Life	10	15	12
PGL	-2,915	26	34
NSG	2,946	-13,405	495

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

Table 5-10. Doubled Counted Savings (Therms) from GPY6

	HEJ	HEReb	MF
Measure Life	10	15	12
PGL	-2,479	8,406	63
NSG	1,902	-30,077	495

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.

Table 5-11. Doubled Counted Savings (Therms) from CY2018

	HEJ	HEReb*	MF	IHWAP	SF
Measure Life	10	15	12	19	20
PGL	-2,479	8,406	63	N/A	N/A
NSG	1,902	-30,077	495	636	-1,606

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.
*This includes the weatherization measures which were considered separately in the CY2018 report but which were part of the HEReb program.

6. APPENDIX 2. TOTAL RESOURCE COST DETAIL

Table 6-1, the Total Resource Cost table for PGL and NSG, includes cost-effectiveness analysis inputs available at the time of finalizing the CY2019 HER impact evaluation report. Additional required cost data (e.g., measure costs, program level incentive and non-incentive costs) are not included in this table and will be provided to evaluation later. Detail in this table (e.g., EULs), other than final CY2019 savings and program data, are subject to change and are not final.

Table 6-1. Total Resource Cost Savings Summary for PGL and NSG

Savings Category	PGL	NSG	NSG Sept Wave
Number of Participants	80,649	59,203	17,933
Effective Useful Life (Years)	5	5	5
Ex Ante Savings, therms	688,083	456,028	
Verified Net Savings Before Uplift Adjust., therms	825,429	713,560	0
Verified Net Savings After Uplift Adjust., therms	772,269	697,446	0

Source: Guidehouse analysis of PGL and NSG program tracking and customer billing data.