NAVIGANT

150 N Riverside Plaza
Suite 2100
Chicago, IL 60606
www.navigant.com

To:	Vince Gutierrez (ComEd); Randy Opdyke (Nicor Gas)
CC:	Jennifer Morris (ICC Staff)
From:	Will Sierzchula and Carly Olig (Navigant)
Date:	March 29, 2019
Re:	Connected Savings CY2018 Heating Season Pre-Period Data Analysis

INTRODUCTION

This memo identifies how additional pre-period data affected energy savings and randomization validation for the joint Nicor Gas and Commonwealth Edison (ComEd) CY2018 Connected Savings thermostat optimization program. CY2018 covered 2018-01-01 through 2018-12-31, but this program evaluation specifically focused on the heating season from November 2017 through April 2018.

Navigant provided CY2018 Connected Savings heating season program savings in an evaluation report. The initial impact analysis only had valid pre-period data from 2017-11-11 through 2017-11-30. Due to this limited data availability and some customers not having any pre-period data, Navigant used a regression model with only post-period data.

After Navigant presented the impact evaluation report, Whisker Labs (the implementer) was able to provide additional pre-period data. The clients Nicor Gas and ComEd requested Navigant determine the extent to which this new data affected program savings. In addition, the data also allowed Navigant to validate randomization (i.e., whether participants and controls were randomly distributed according to usage).

The following summarizes Navigant's key findings from the heating season analysis incorporating the expanded pre-period data set.

Finding 1. Approximately 100 participants and 80 controls (representing 8.5% of all accounts) did not have observations in the expanded pre-period data set.

Finding 2. Including the additional pre-period data resulted in a decrease in the savings estimate standard errors. However, only one of the pre-post models was statistically significant. These results suggest that pre-post models with slightly larger sample sizes are likely to result in statistically significant savings.

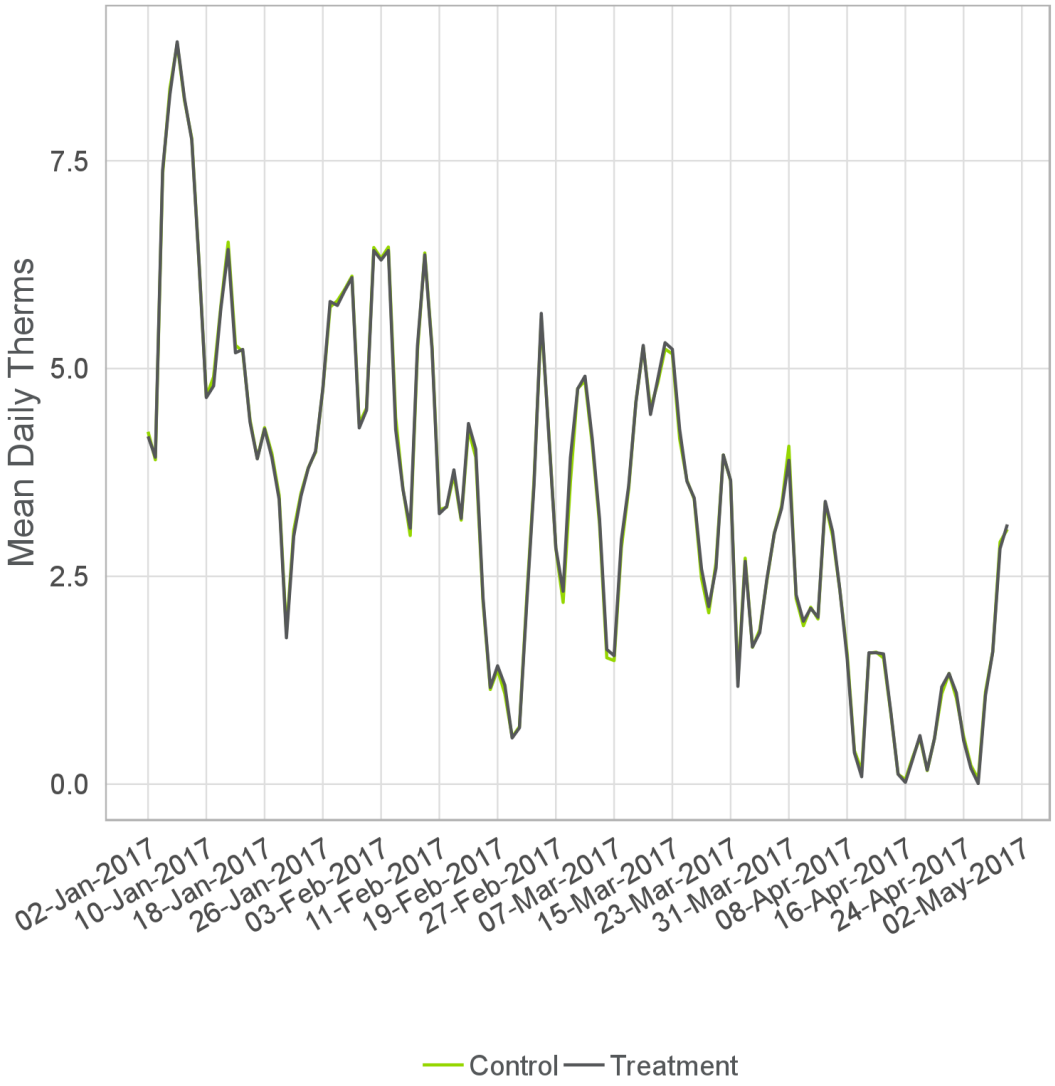
Finding 3. Based on their pre-period usage, customers were randomly assigned to either participant or control groups, validating the RCT.

Recommendation 1. Since the RCT validation showed balanced participant and control groups, an analysis of post-only data should be unbiased. Consequently, Navigant recommends using the post-only results (2.85% savings) because they include the 8.5% of accounts that did not have any pre-period data.

RCT VALIDATION

To test whether Connected Savings accounts were randomly assigned to participant or control groups, Navigant visually compared gas consumption during the pre-period, and also ran a regression on pre-period usage with the treatment indicator as the independent variable. Figure 1 illustrates almost identical participant and control usage during the pre-period.

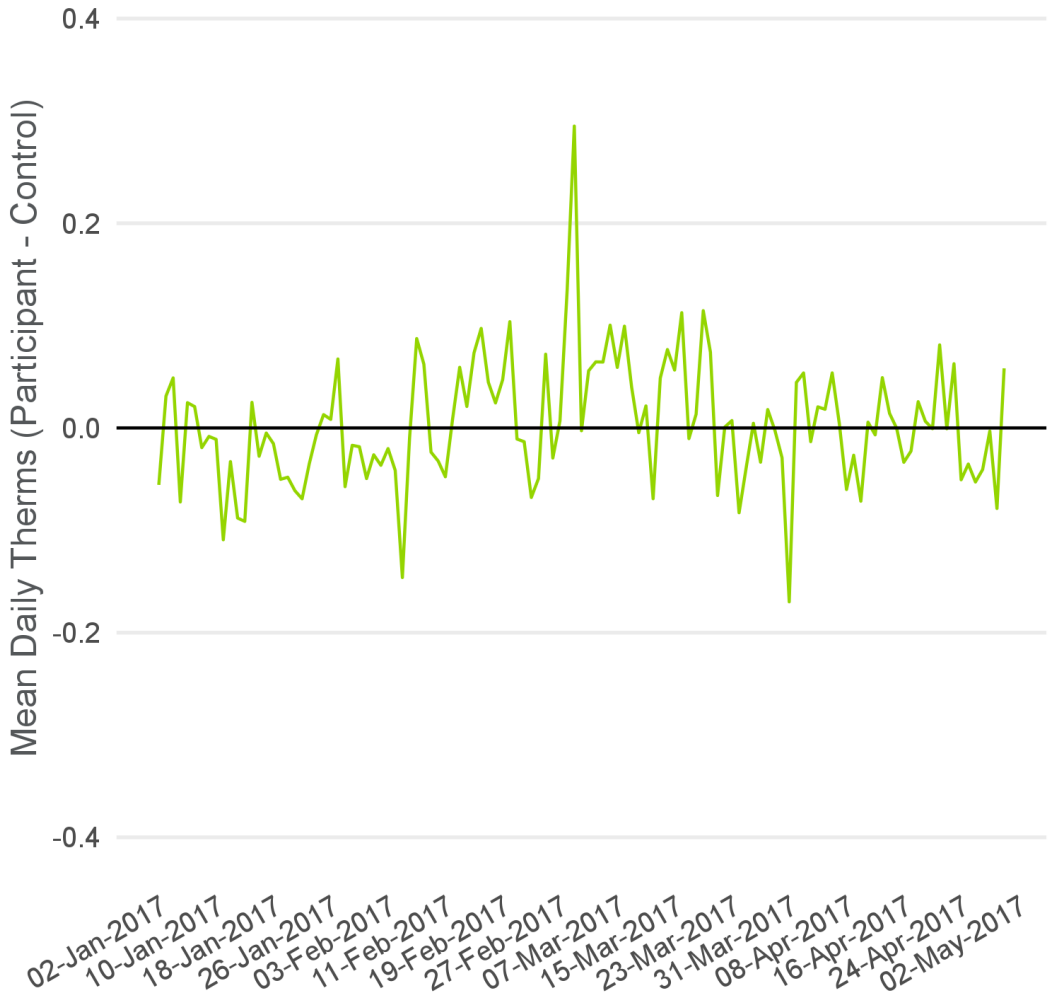
Figure 1. Pre-Period Usage Comparison



Source: Whisker Labs telemetry data and Navigant team analysis.

Figure 2 shows the difference in average daily therm usage between participants and controls during the pre-period. The plot shows that neither group had consistently higher or lower pre-period usage, validating program randomization.

Figure 2. RCT Validation – Participant Minus Control Usage in Pre-Period



Source: Whisker Labs telemetry data and Navigant team analysis.

In addition, Table 1 provides regression results that show the treatment variable was not statistically significant in describing participant and control usage during the pre-period. These results validate participant and control randomization.

Table 1. RCT Validation Regression Results

	Estimate	Std. Error	T Statistic	P Value
(Intercept)	3.60	0.07	53.34	0.00
treatment	-0.02	0.09	-0.16	0.87

Source: Whisker Labs telemetry data and Navigant team analysis.

REGRESSION RESULTS

To determine how the additional pre-period data affected program savings, Navigant combined it with the post-period data and ran several regressions to test sensitivity. Navigant conducted the same data cleaning steps on both the new pre-period data and the existing post-period data. Table 2 provides the new pre-period data cleaning results, while the evaluation report's data cleaning results are available in Table 3.

Comparing Table 2 and Table 3 shows fewer participants and controls had data in the pre-period. Further examination of these records identified approximately 100 participants and 80 controls did not have pre-period data. However, the data cleaning steps dropped a similar number of customers and observations across both participants and controls, suggesting the data remained balanced for regression analysis.

Table 2. Pre-Period Data Cleaning

Data Cleaning Step	Customers		Observations		Customer % Drop		Observation % Drop	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
Raw interval data	955	963	9,456,835	9,435,867				
Missing combustible heat interval	955	963	9,456,835	9,435,867	0%	0%	0%	0%
Aggregate to daily	955	963	105,016	105,881	0%	0%	99%	99%
Remove days non-combustible runtime	955	963	101,317	101,710	0%	0%	4%	4%
Filter out incomplete days	951	957	83,285	81,815	0%	1%	18%	20%

Source: Whisker Labs telemetry data and Navigant team analysis.

Table 3. Evaluation Report Data Cleaning

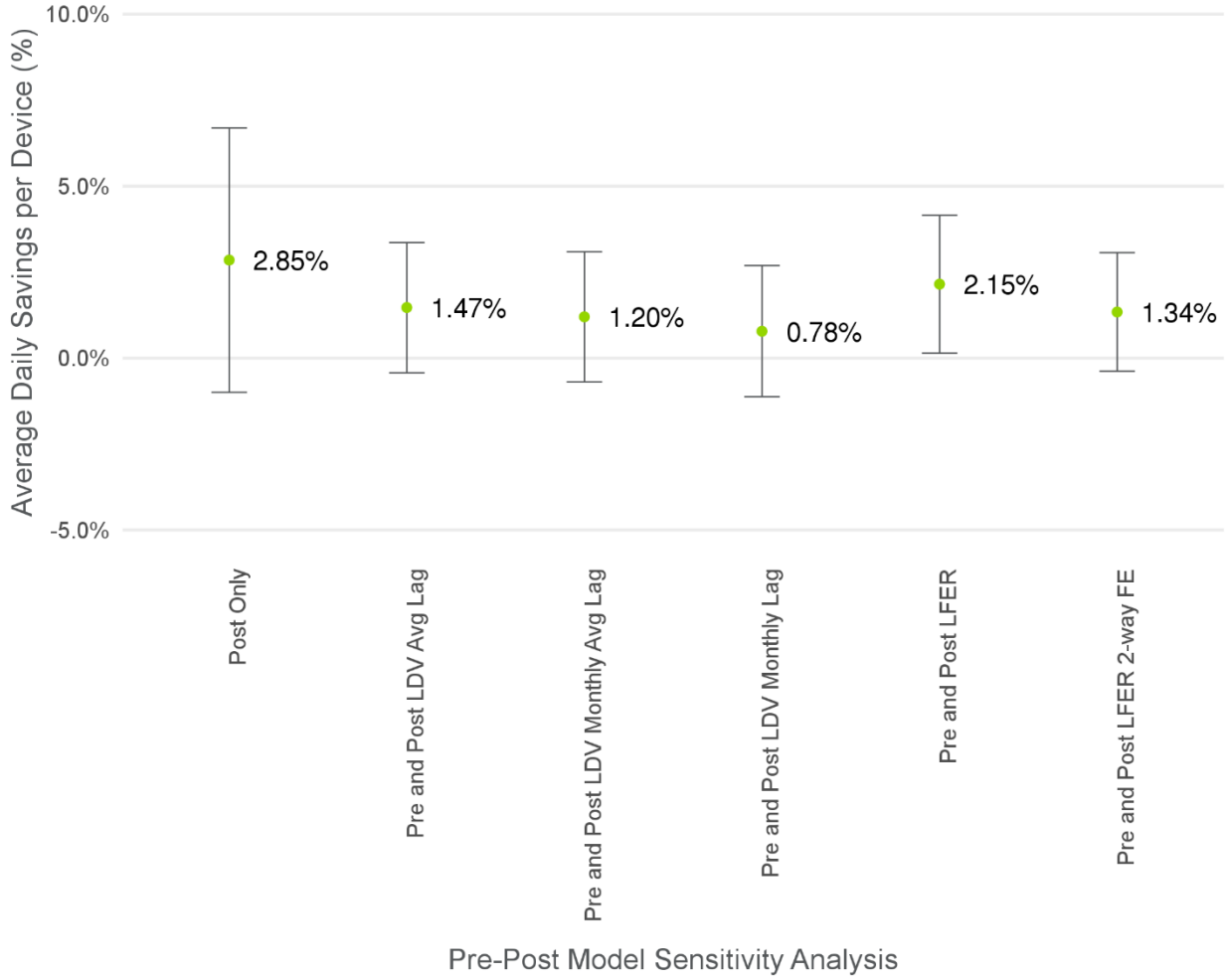
Data Cleaning Step	Customers		Observations		Customer % Drop		Observation % Drop	
	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Control
Raw interval data	1,081	1,061	27,285,569	27,235,062				
Missing combustible heat interval	1,049	1,034	15,153,109	15,300,274	2.4%	2.1%	10.2%	8.4%
Aggregate to daily	1,049	1,034	167,033	170,353	0.0%	0.0%	98.9%	98.9%
Remove days non-combustible runtime	1,048	1,034	164,601	167,246	0.1%	0.0%	1.5%	1.8%
Filter out incomplete days	1,045	1,031	137,790	135,877	0.3%	0.3%	16.3%	18.8%

Source: Whisker Labs telemetry data and Navigant team analysis.

After combining the cleaned pre- and post-period data sets, Navigant ran several regression models to test the sensitivity of savings and standard errors. The model specifications are shown in Table 4 in the Appendix of this memo. Figure 3 shows the results of these sensitivity analyses. The original post only model is shown first and then five pre and post sensitivity models. The savings estimate varies from 0.78% to 2.85% across the models but none of these estimates are statistically different from one another. Only one of the models (the Pre and Post Linear Fixed Effects Regression [LFER]) is statistically

different from zero (i.e., the confidence bound does not cross the 0% savings line). These results suggest that models with pre and post-period data and a slightly larger treatment group would likely result in statistically significant savings.

Figure 3. Sensitivity Analyses Percent Savings



Source: Whisker Labs telemetry data and Navigant team analysis.

APPENDIX

The model specifications for each of the models presented in Figure 3 are shown in Table 4.

Table 4. Model Specifications

Model Name	Specification
Post Only	$EDU_{it} = \beta_1 Treat_i + \gamma_m + \varepsilon_{it}$
Pre and Post LDV Avg Lag	$EDU_{it} = \beta_1 Treat_i + \gamma_m + \beta_2 LDV\_Avg_i + \varepsilon_{it}$
Pre and Post LDV Monthly Avg Lag	$EDU_{it} = \beta_1 Treat_i + \gamma_m + \gamma_m \cdot LDV\_Avg_i + \varepsilon_{it}$
Pre and Post LDV Monthly Lag	$EDU_{it} = \beta_1 Treat_i + \gamma_m + \gamma_m \cdot LDV_{im} + \varepsilon_{it}$
Pre and Post LFER	$EDU_{it} = \beta_1 Treat_i \cdot post_t + \beta_2 post_t + \alpha_i + \varepsilon_{it}$
Pre and Post LFER 2-way FE	$EDU_{it} = \beta_1 Treat_i \cdot post_t + \gamma_m + \alpha_i + \varepsilon_{it}$

Source: Navigant

Where:

EDU_{it}	is estimated daily usage of therms by device *i* on day *t*
$Treat_i$	is a binary variable taking a value of 1 when device *i* is in the treatment group and 0 otherwise
γ_m	is a time-specific fixed effect for month *m;* this picks up temporal differences across months, like weather and daylight hours
LDV_Avg_i	is average lagged estimated daily usage of therms by device *i* across all pre-period data available for device *i*
LDV_{im}	is average lagged estimated daily usage by device *i* across all pre-period observations in the same month *m* of the pre-period as day *t* is in in the post-period
$post_t$	is a binary variable taking a value of 1 when day *t* is in the post-period and 0 otherwise
α_i	is a device-specific fixed effect for device *i*; this captures all attributes about device *i* that do not vary over time
ε_{it}	is the cluster-robust error term for device *i* during day *t;* cluster-robust errors account for heteroskedasticity and autocorrelation at the household level