NAVIGANT

30 South Wacker Drive
Suite No. 3100
Chicago, Illinois 60606

Draft

To: Jim Jerozal, John Madziarczyk, Bridgid Lutz, Steve Grzenia, Nicor Gas; Scott Dimetrosky, Apex Analytics; Ted Weaver, First Tracks Consulting; Vincent Gutierrez, ComEd; Patrick Michalkiewicz, Peoples Gas; Susan Nathan, Applied Energy Group; Paige Knutsen, Laura Pavlot, Franklin Energy; Jennifer Hinman Morris, David Brightwell, ICC Staff; Celia Johnson, Future Energy Enterprises

From: Christy Zook and Chelsea Lamar, Navigant

CC: Randy Gunn, Charley Budd, Jeff Erickson, Laura Agapay-Read, Kevin Grabner, Rob Neumann, Josh Arnold, Katherine Wolf, Meghan Sposato, Navigant

Date: February 16, 2016

Re: GPY4/EPY7 Net-to-Gross Ratio Estimates for Future Use for the Nicor Gas, ComEd, Peoples Gas, and North Shore Gas Elementary Energy Education Program

This memo presents results from Navigant's GPY4/EPY7 evaluation activities that will support our January 7, 2016 delivery of net-to-gross (NTG) values that will be used prospectively in GPY6/EPY9 for the Elementary Energy Education (EEE) program[1]. Navigant calculated net-to-gross values using two algorithms: one from the draft Illinois TRM statewide approach[2] and the other from the approach Navigant used in GPY1/EPY4. We will provide additional results from our GPY4/EPY7 evaluation in separate evaluation reports for each utility.

ELEMENTARY ENERGY EDUCATION

In GPY4/EPY7, the EEE program was jointly offered by Nicor Gas, ComEd, Peoples Gas (PG), and North Shore Gas (NSG). The program's primary focus is to produce natural gas and electricity savings in the residential sector by motivating 5th grade students and their families to reduce energy consumption from water heating and lighting in their home. Students take home a free energy saving kit that includes high efficiency showerheads, bathroom and kitchen faucet aerators, and CFLs (only in kits in the ComEd service territory).

NET-TO-GROSS RATIO ESTIMATES

The evaluation team's net-to-gross estimates using the draft Illinois TRM approach (TRM), as well as the GPY1/EPY4 approach (historic) for the program, are shown in Table 1 below. The two approaches

[1] This memo was originally delivered December 18, 2015 and was finalized February 16, 2016.
[2] IL-TRM_Attach A_IL-NTG Methods_10_02_15_DRAFT.docx

produce very similar free ridership results for electric measures on the whole (0.36 vs. 0.34) but the gas measures NTG differs by 0.1 (0.27 TRM, 0.17 historic). This result is mainly driven by a lower TRM CFL NTG value (only electric) and higher TRM NTG values for the other measures (both electric and gas).

Table 1. Program Net-to-Gross Ratio and Components from Two Approaches

	Free Ridership		Spillover		NTG	
	TRM	GPY1/EPY4	TRM	GPY1/EPY4	TRM	GPY1/EPY4
Showerheads	0.29	0.19	0.11	0.14	0.82	0.95
Bathroom Faucet Aerators	0.20	0.13	0.12	0.15	0.92	1.01
Kitchen Faucet Aerator	0.23	0.13	0.12	0.14	0.89	1.01
CFL	0.51	0.62	0.18	0.10	0.67	0.48
Unlike SO, Electric and Gas			0.02	-		
Electric Measures	**0.36**	**0.34**	**0.16**[3]	**0.12**	**0.80**	**0.78**
Gas Measures	**0.27**	**0.17**	**0.13**[4]	**0.14**	**0.87**	**0.97**

The evaluation team also conducted a free ridership (FR) sensitivity analysis where the evaluators tested an alternative method for combining the non-program, timing, and quantity scores, to report on the sensitivity of results to these changes. The sensitivity analysis only applied to measures that included a quantity component: the CFLs and the bathroom faucet aerators. The results of the alternate FR algorithm can be seen in Table 2 below. The alternative FR methodology resulted in a slightly higher FR for CFLs and no change to the bathroom faucet aerators FR rates.

Table 2. Free-Ridership Estimates Compared to Alternative Method

Measure	Draft IL TRM FR	Alternative FR
CFL	0.51	0.54
Bathroom Faucet Aerator	0.20	0.20

Source: Evaluation Analysis

[3] This represents unlike SO added to the weighted average of the electric measure level like SO (0.14).
[4] This represents unlike SO added to the weighted average of the gas measure level like SO (0.11).

DATA COLLECTION FOR NET TO GROSS ESTIMATES

Table 3 below summarizes primary data sources that Navigant used to estimate the NTGR for the program. The survey achieved 5.9 percent precision at a 90 percent confidence interval.

Table 3. Primary Data Sources

Method	Subject	Combined Target Completes	Combined Actual Completes	Completed	Confidence Precision
Take-Home Survey	GPY4/EPY7 Program Participants	258[5]	191	May 15, 2015	90/6

Source: Evaluation Analysis

TRM AND HISTORIC NET-TO-GROSS METHODOLOGIES

As part of the GPY4/EPY7 NTG analysis, the evaluation team calculated NTG using two methods, the draft Illinois TRM NTG methodology and the GPY1/EPY4 NTG methodology. This was done so that the NTGR for the different program years can be compared using the same algorithm. This section describes the free-ridership and spillover methodologies that were used in the draft Illinois TRM approach as well as in the GPY1/EPY4 approach.

The free-ridership and spillover rates were assessed using the same self-reported data gathered through Navigant's participant survey. The participant survey included questions to identify installations that might have occurred if the utilities had not funded the EEE program. This data allows Navigant to estimate free-rider ratios—a factor that effectively deducts "free-riders" from the gross savings identified via the impact analysis. The survey also included questions to help identify participant spillover effects.

The final NTGRs for each measure are calculated as:

$$NTG = 1 - [Free\ Ridership] + [Spillover]$$

Where,
Free ridership is the energy savings that would have occurred even in the absence of program activities and sponsorship, expressed as a percent of gross impact.

And,
Spillover is the energy savings that occurred as a result of program activities and sponsorships, but was not included in the gross impact accounting, expressed as a percent of gross impact.

[5] The sample goal was designed to reach statistical significance for each utility territory

Free Ridership – Draft Illinois TRM Approach

Free ridership cannot be measured directly due to absent empirical data regarding the counterfactual situation. Thus, free ridership is assessed as a probability score for each measure. The evaluation relies on self-reported data collected during participant paper-based surveys to assign free ridership probability scores to each measure. More specifically, for each measure, the following questions were posed to each measure recipient[6]:

> FR1. On a scale of 0 to 10, with 0 meaning "No, I was not planning to buy this high efficiency item" and 10 meaning "Yes, I was planning to buy this high efficiency item." Were you planning to buy the same items in the kit before you received the kit?
>
> FR2. When were you planning to purchase and install them?

For measures with a quantity of greater than one, the following question was also included:

> FR3. Were you planning to purchase the same number of [measures] as in the kit on your own?

The following question was also asked of all participants and used as a consistency check:

> CC1. Before you received the [measure] in the kit, was your family already planning to purchase the same high efficiency [measure] from the store?

Free Ridership Scoring—TRM Approach

The free ridership data was assembled into a probability score in a step-by-step fashion, applying the following logic:

> If the participant indicated a low likelihood that they had been planning to purchase the item before receiving it in their kit (FR1 <= 3), the participant's response to FR1 divided by 10 is considered the participant's free ridership score. In the IL TRM, the response to FR1 is referred to as the "Non-Program Score".
>
> If the participant gave a response to FR1 greater than 3, the timing score (FR2) and the quantity score (FR3, where applicable) were first averaged, and then the response to FR1 was averaged with the average of the timing and quantity, if the timing and quantity score are less than the response to FR1.
>
> The timing score is 0.5 if the high efficiency measure would have been purchased within 6 months, 0.25 if it would have been purchased within 6 months to a year later, and 0 if it would have been purchased more than a year later.

[6] The survey instrument instructions directed an adult to complete the survey.

The corresponding formula for calculating free ridership is shown below:

$$if\ FR1 > 3\ and\ FR2, FR3 < \frac{FR1}{10},$$
$$then\ FR = Average\left[\frac{FR1}{10}, Average(FR2, FR3)\right],$$

$$else\ FR = \ FR1/10$$

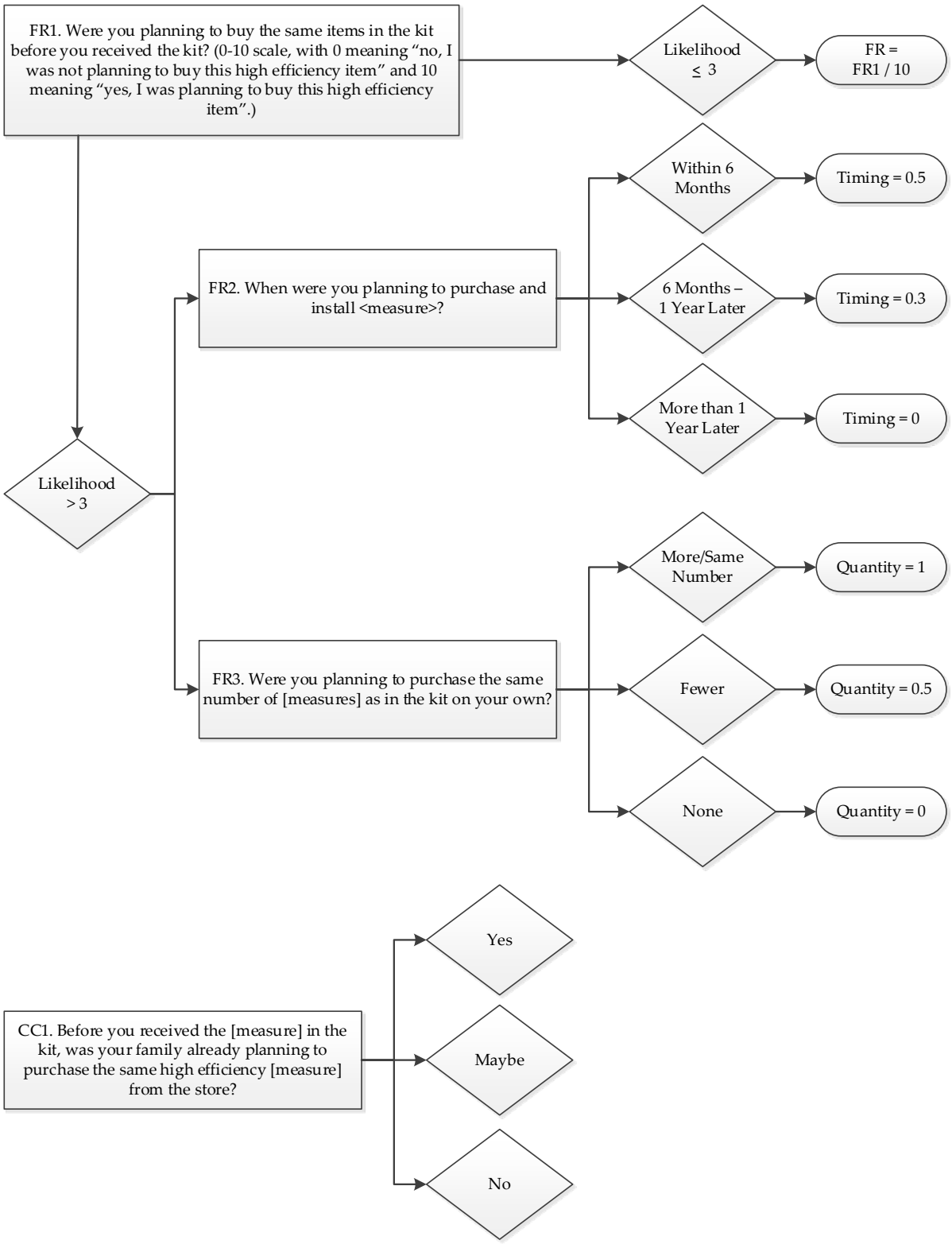
Note that in the above formula, if FR1 is invalid (missing or "don't know"), then the participant's responses for NTG determination are disqualified. Eight participants were removed from the analysis based on their response to FR1. Participants were only removed from the FR calculations for the individual measure(s) where they had an invalid response for FR1 but were included for the other measures.

If a participant 1) replied to the consistency check (CC1) that they were planning on purchasing the measure before they received their kit (a "yes" response, indicating high or full free ridership) and 2) had a calculated FR of less than 0.5, they were removed from the analysis because their responses are not consistent. Likewise, if a participant 1) indicated that they were not planning on purchasing the measure (a "no" response to CC1, indicating low or no free ridership) and 2) had a calculated FR rate of greater than 0.5, they were also removed from the analysis. Participants who responded "maybe" to CC1 were not included in the consistency check. Twenty participants were removed from the analysis based on their FR rates and responses to the consistency check question. Participants were only removed from the FR calculations for the individual measure(s) where they failed the consistency check but were included for the other measures.

This approach is a modification of that used in the Nicor Gas Rider 29 evaluation to add precision and to approximate the free ridership approaches currently proposed by the Illinois TRM working group. The free ridership methodology is presented in Figure 1 below.

The free-ridership rate was calculated for each individual kit component and participant. The individual free-ridership rates were then averaged to calculate the free-ridership rate per component, and weighted by individual savings, for measures where the quantity is greater than one. The program free-ridership rate was calculated using a weighted average by component savings. The component savings were calculated using Illinois TRM deemed values and the specific component values, where appropriate. The free-ridership rates were then weighted by program savings in order to calculated overall free-ridership for each fuel type (gas or electric).

Figure 1. Participant Free-Ridership Algorithm—TRM Approach



Free Ridership - GPY1/EPY4 Historic Approach

The GPY1/EPY4 FR methodology used the same questions as the draft Illinois TRM NTG methodology, with the inclusion of CC1 as part of the algorithm.

> CC1. Before you received the [measure] in the kit, was your family already planning to purchase the same high efficiency [measure] from the store?
>
> FR1. On a scale of 0 to 10, with 0 meaning "No, I was not planning to buy this high efficiency item" and 10 meaning "Yes, I was planning to buy this high efficiency item." Were you planning to buy the same items in the kit before you received the kit?
>
> FR2. When were you planning to purchase and install them?

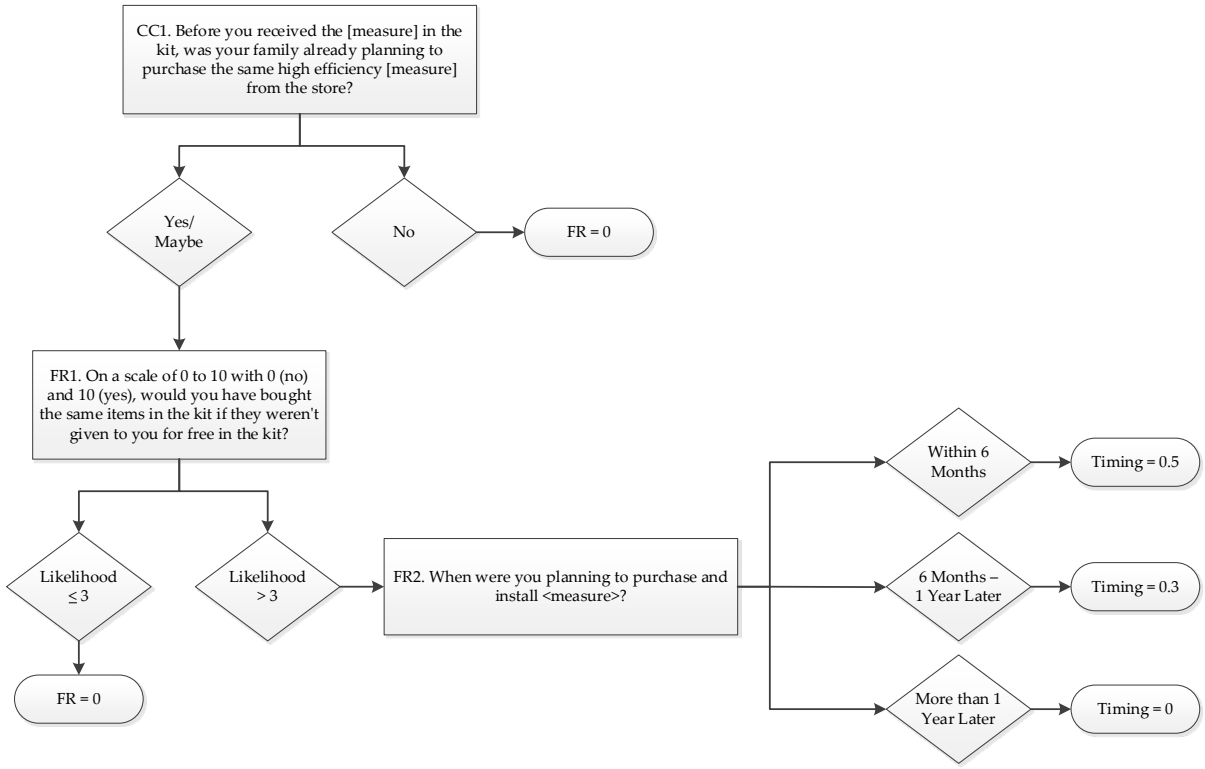
Free Ridership Scoring—Historic Approach

The free ridership data was assembled into a probability score in a step-by-step fashion, applying the following logic:

> If the participant reported that they were not planning on purchasing the measure before they received their kit, then the probability of free ridership for that participant is estimated to be zero (based on CC1 above). Similarly, if the participant reported likelihood of purchasing the same measures as provided in the kit less than or equal to 3 (on a 0-10 scale), then the probability of free ridership is estimated to be zero (based on the response to FR1). If neither of the above criteria holds, then responses to question FR2 (the timing score) and FR1, likelihood of purchasing the measures in the absence of the program (the non-program score), were averaged and divided by 10 to calculate the probability of free ridership. The corresponding formula for calculating free ridership is shown below:
>
> $$if\ CC1 = "No"\ or\ FR1 \leq 3, then\ FR = 0,$$
>
> $$else\ FR = \ Average(\frac{FR1}{10}, FR2)$$
>
> Note that in the above formula, if CC1 is invalid (missing or "don't know") then the participant's responses for NTG determination are disqualified.

The free-ridership rate was calculated for each individual kit component and participant. The individual free-ridership rates were then averaged to calculate the free-ridership rate per component and weighted by individual savings, for measures where the quantity is greater than one. The program free-ridership rate was calculated using a weighted average by component savings. The component savings were calculated using Illinois TRM deemed values and the specific component values, where appropriate. The free-ridership rates were then weighted by program savings in order to calculate overall free-ridership for each fuel type (gas or electric).

Figure 2. GPY1/EPY4 Participant Free-Ridership Algorithm—Historic Approach



Free Ridership Sensitivity Analysis of Historic Approach

In addition to reporting results based on the above algorithms, the evaluators tested an alternative method for combining the non-program (FR1), timing (FR2), and quantity scores (FR3), to report on the sensitivity of results to these changes. This information is intended to inform the TRM NTG algorithm development process. The primary difference between the draft IL TRM FR method and the alternative method is how the non-program, timing, and quantity responses are averaged. In the draft IL TRM FR method, the timing and quantity responses are first averaged, then that number is average with the non-program score (divided by 10). In the alternative method, the non-program score (divided by 10), timing score, and quantity score are averaged together. The free ridership alternative method was calculated using the equation below:

$$if\ FR1 > 3\ and\ FR2, FR3 < \frac{FR1}{10},$$
$$then\ FR = Average\left[\frac{FR1}{10}, FR2, FR3\right],$$
$$else\ FR = \ FR1/10$$

where the timing and quantity scores were assigned as they were in the draft Illinois TRM FR analysis.

The sensitivity analysis only applied to measures that included a quantity component: the CFLs and the bathroom faucet aerators. Because there was only one high efficiency showerhead and one kitchen faucet aerator in the each kit, the free-ridership calculations for these measures did not include the quantity score, and therefore the sensitivity analysis could not be performed on them. The results of the alternate FR algorithm can be seen below. The alternative FR methodology resulted in a slightly higher FR for CFLs and no change to the bathroom faucet aerators FR rates.

Table 4. Free-Ridership Estimates Compared to Alternative Methods

Measure	Draft IL TRM FR	Alternative FR
CFL	0.51	0.54
Bathroom Faucet Aerator	0.20	0.20

Source: Evaluation Analysis

Spillover – Draft Illinois TRM Approach

The objective of the spillover assessment is to estimate the impact arising from efficient measures installed as a result of the program that were not incented by the program. The evaluation relied on self-reported data collected during the paper-based participant survey to identify these measures and assess the role of the program in the decision to install. The spillover methodology approximates the spillover methodology currently proposed by the Illinois TRM working group.[7] Like spillover (spillover from program measures) and unlike spillover (spillover from other efficient measures) were estimated and are defined below.

Like Spillover—TRM Approach

For each measure installed through the program, the following questions are posed to each measure recipient:

> SP1. AFTER the program came to your school, did you BUY and INSTALL any showerheads, faucet aerators, or CFLs like the ones in the kit?
>
> SP2. Please note how many you bought and installed.
>
> SP3. Did you receive a rebate from your gas or electric utility for your purchase?
>
> SP4. If you bought more showerheads, aerators, or CFLs after the program, how likely was it that you bought them because of your experience with the kit? (0-10 scale)

[7] IL-TRM_Attach A_IL-NTG Methods_10_02_15_DRAFT.docx

Unlike Spillover—TRM Approach

A similar series of questions were asked to participants regarding unlike spillover:

> USP1: Did you complete any additional energy efficiency upgrades after receiving the kit (for example, purchase LED bulbs, weatherize your home, or purchase a high efficiency appliance)?
>
> USP2: Did you receive an incentive from your gas or electric utility for your upgrade?
>
> USP3: If you completed energy efficiency upgrades after receiving the kit, how likely was it that you bought them because of the kit?

Spillover Scoring—TRM Approach

The survey data was assembled into an assessment of spillover impact through application of the following method:

> If the customer installed additional units of the measure following their participation, did not receive an incentive from their gas or electric utility for the upgrade, and the program was highly influential in the decision to install those measures, the adoption is considered to be potentially program spillover:
>
> *If SP1=Yes, SP3=No, and SP4 >7,*
>
> *then SO = (SP2*Measure Savings) / Program Measure Savings*

Any savings associated with spillover were weighted against the total savings of the participant sample for the particular measure to establish a measure-specific spillover rate. The spillover methodology is shown in Figure 3 below. The spillover rate was calculated for each individual kit component and participant. The individual spillover rates were then averaged to calculate the spillover rate per component. The spillover rate by fuel type (gas or electric) was calculated using a weighted average by component savings. The component savings were calculated using Illinois TRM deemed values and the specific component values, where appropriate. The participants with spillover had an assigned spillover value based on their influence score, and the fraction of the measure savings out of the total program measure savings.

Unlike Spillover Scoring—TRM Approach

If the customer completed additional energy efficiency upgrades after receiving the kit, did not receive an incentive from their gas or electric utility for the upgrade, and reasoning for completing these upgrades was somewhat related to the customer's experience with the kit, their savings contributed to unlike spillover as calculated below:

If USP1=Yes, USP2=No, and USP3 >7,

then unlike SO = $[\sum Estimated\ Energy\ Savings]$*/ Total Sample Savings*

These percentages were calculated separately for participants using gas and electric heat, with estimated energy savings in units of therms or kWh respectively.

Figure 3. Participant Like Spillover Algorithm—TRM Approach

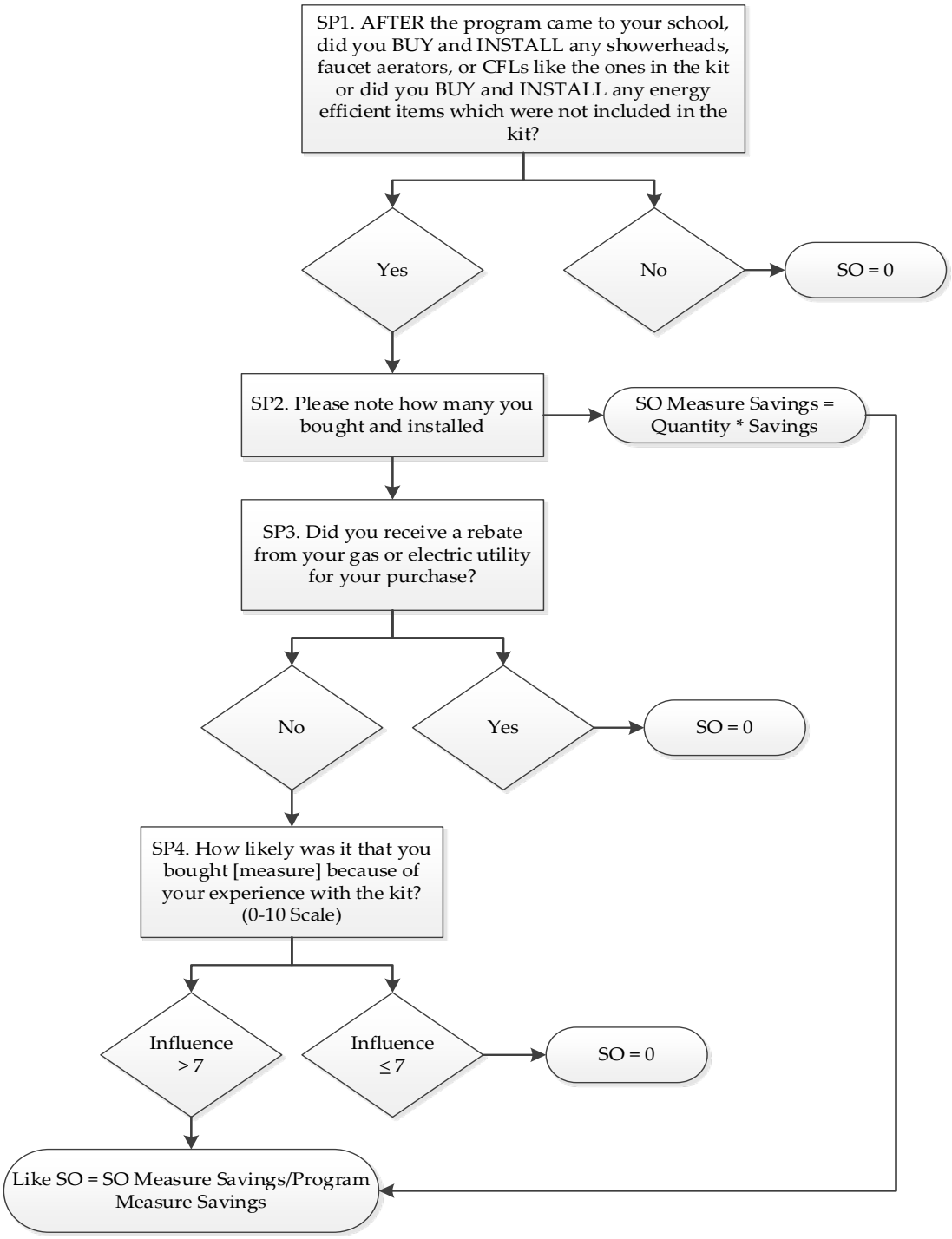
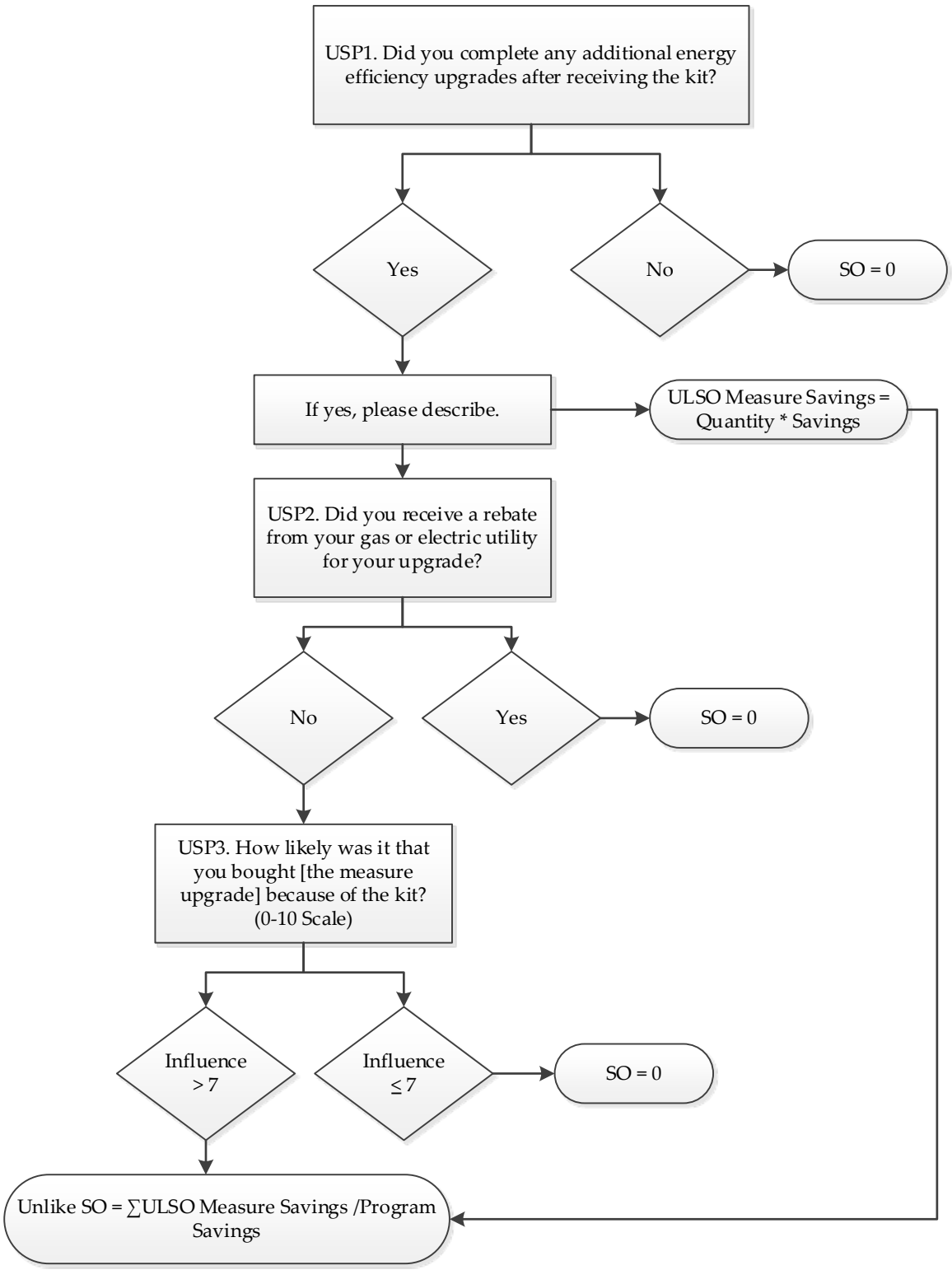


Figure 4. Participant Unlike Spillover Algorithm—TRM Approach



The estimations of both like and unlike spillover by measure are presented in Table 5 and Table 6 below.

Table 5. Like Spillover by Measure Type—TRM Approach

Measure	Measure-level Spillover
Showerhead	0.11
Bathroom Aerator	0.12
Kitchen Aerator	0.12
CFL	0.18

Source: Evaluation Analysis

Table 6. Unlike Spillover Estimates by Heating Type—TRM Approach

	Gas	Electric
Unlike SO	0.02	0.02

Source: Evaluation Analysis

Spillover – GPY1/EPY4 Historic Approach

The objective of the spillover assessment is to estimate the impact arising from efficient measures installed as a result of the program that were not incented by the program. The evaluation relied on self-reported data collected during the paper-based participant survey to identify these measures and assess the role of the program in the decision to install.

For each measure installed through the program, the following questions are posed to each measure recipient:

> SP1. AFTER the program came to your school, did you BUY and INSTALL any showerheads, faucet aerators, or CFLs like the ones in the kit?
> SP2. How many additional measures did you install?
> SP3. If you bought more showerheads, aerators, or CFLs after the program, how likely was it that you bought them because of the program? (0-10 scale)

Spillover Scoring—Historic Approach

The survey data was assembled into an assessment of spillover impact through application of the following method:

> If the customer installed additional units of the measure following their participation, and the program was highly influential in the decision to install those measures, the adoption is considered to be potentially program spillover:
>
> *[If SP1=1 and SP3 is greater than 7, then adoption is spillover]*

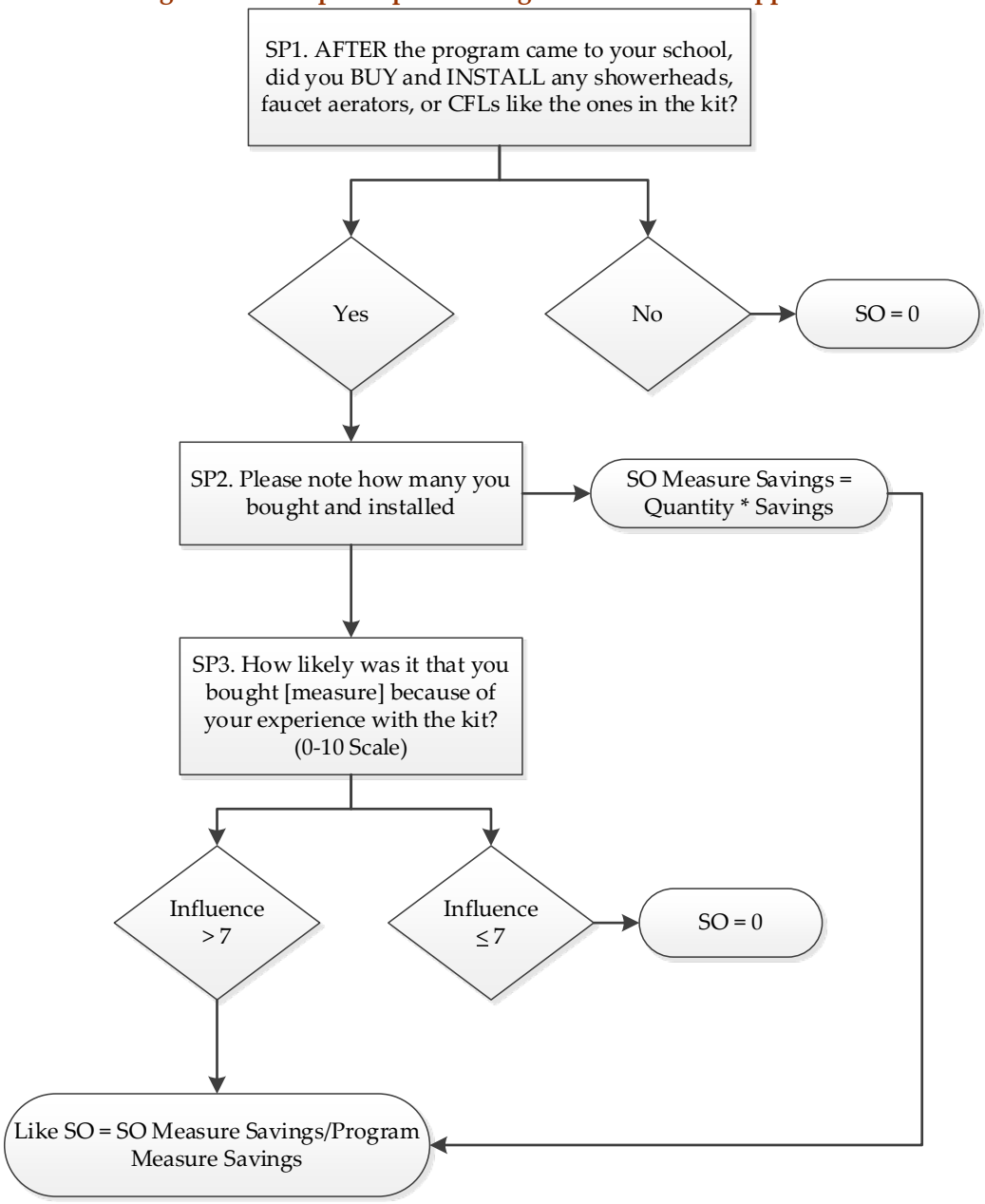
Any savings associated with spillover were weighted against the total savings of the participant sample for the particular measure to establish a measure-specific spillover rate. The spillover methodology is shown in Figure 5 below. The spillover rate was calculated for each individual kit component and participant. The individual spillover rates were then averaged to calculate the spillover rate per component. The spillover rate by fuel type (gas or electric) was calculated using a weighted average by component savings. The component savings were calculated using Illinois TRM deemed values and the specific component values, where appropriate. The participants with spillover had an assigned spillover value based on their influence score, and the fraction of the measure savings out of the total program measure savings.

CFL-Specific Adjustments to Spillover—Historic Approach

The impact credit granted for CFL spillover adoptions must avoid double counting the impact credit accrued already through the ComEd midstream residential lighting program. Navigant uses the approach established in the ComEd Single Family PY3 evaluation that assumes that 1) the market share of program bulbs is not a readily available number and 2) the residential lighting program PY3 evaluation results indicated a substantial amount of free ridership (41percent), and there is no reason that one program's free ridership cannot be another program's net impact. Thus, it is not necessary that bulbs be un-incented for them to legitimately qualify for credit under the Single Family Program.[8] Due to the uncertainty in this area, the evaluation team takes the conservative approach used in the PY3 Single Family evaluation and assumes that only 50 percent of the impact arising from CFL spillover adoptions is creditable to the program. Again, even if these customers purchased a discounted bulb, the purchase decision was either influenced by both programs (making the 50 percent assumption reasonable) or influenced by only the EEE program (making the 50 percent assumption conservative).

[8] There is some available evidence regarding the CFL market share of residential lighting program bulbs. The PY3 residential lighting general population survey revealed that 87 percent of CFLs are purchased at stores participating in the ComEd lighting program. Among program stores, the shelf space dedicated to ComEd program CFL bulbs is 53 percent of the overall shelf space dedicated to CFLs (for standard bulbs), and 62 percent for specialty bulbs. If we assume shelf space relates directly to sales share, then 46 percent of standard CFLs and 54 percent of specialty bulbs are Residential Lighting program bulbs.

Figure 5. Participant Spillover Algorithm—Historic Approach



Navigant conducted a paper survey with a stratified random sample with a goal of 258 participating customers from GPY4/EPY7. The actual number of surveys returned from participating customers was 191 providing a 6 percent precision at a 90 percent confidence interval at the program level.

APPENDIX

Joint Utility
EEE-Super Savers Pro