

#### Memorandum

**To:** Illinois SAG

**From:** The Opinion Dynamics Evaluation Team

**Date:** August 21, 2023

**Re:** Deviation from IL-TRM Midstream Free-Ridership Protocol for 2023 Research

This memo outlines a proposed deviation from the Illinois Technical Reference Manual (IL-TRM) Midstream Free-Ridership Protocol (“Midstream Protocol”) to be used by Opinion Dynamics for 2023 evaluation research.

This deviation will apply to two midstream program free-ridership research efforts currently underway:

* Interviews with participating distributors in Ameren Illinois Company (AIC)’s Business Midstream Initiative – HVAC channel and Residential Single Family Market Rate Initiative – Midstream HVAC channel
* Interviews with participating distributors in Ameren Illinois Company (AIC)’s Business Midstream Initiative – Lighting channel

These deviations are to ensure that research conducted in summer 2023 is sound and measures the correct concepts.

#### Background

During the past two years, the Illinois Net-to-Gross Working Group (NTG WG) has dedicated significant time to updating the Core Non-Residential Free-Ridership Protocol. As of April 2023, the Illinois independent evaluators (Opinion Dynamics and Guidehouse) reached consensus on substantial revisions to this protocol and issued an interim update to IL-TRM Attachment A (the Illinois Statewide Net-to-Gross Methodologies) dated April 27, 2023, for use in summer 2023 evaluation research. The NTG WG will issue a final version of Attachment A with these updates for inclusion in the final IL-TRM V12.0 this fall.

Multiple other Attachment A protocols are based on the Core Non-Residential Free-Ridership Protocol, including the Midstream Protocol. While making updates to Attachment A, the independent evaluators also revised these protocols to align with the new Core Non-Residential Free-Ridership Protocol. While developing data collection instruments for 2023 evaluation research, Opinion Dynamics concluded that the updated Midstream Protocol does not measure the appropriate concepts and needs to be modified to ensure high-quality evaluation research.

#### Concerns

Our concerns with the Midstream Protocol as outlined in the April 27, 2023 version of Attachment A are twofold:

1. Language in the Protocol focuses partially (but not comprehensively) on “sales strategies” and does not appropriately tie sales strategies to sales volume, leading to FR estimates that may be misaligned with program outcomes.
2. The Protocol’s Counterfactual Free-Ridership Score does not appropriately handle partial effects of program interventions.
	1. For example, using the Protocol’s existing constructs, any distributor who indicated that they would have used some, but not all, of the same sales strategies in the absence of the program would be considered a full non-free-rider – even if the program caused only minor components of the distributor’s sales strategies to change.
	2. Similarly, if the Protocol was refocused on changes in sales volume, rather than changes in sales strategies (as we later propose), distributor indication that the program caused any change in their sales of efficient equipment would lead to treatment of a distributor as a full non-free-rider, even if the program caused only a small change in their total sales volume.

#### Resolution

Our proposed deviation involves two revisions to the questions used in the Midstream Protocol to develop the Program Influence and Counterfactual FR Scores, while combining the scores in the same manner as currently prescribed.

1. To address our first concern, while reminding the respondent of the sales strategies the program may have affected, we will reframe the Program Influence and Counterfactual questions to focus on program-incentivized sales volume rather than sales strategies. For example, the primary Program Influence question will read:

Thinking about your sales over the past year, on a scale of 0 to 10, where 0 means “Not at all influential” and 10 means “Extremely influential”, how influential were the program elements I listed on your sales of program-incentivized equipment?

1. To address our second concern, we will reframe the Counterfactual follow-up question to ask respondents to estimate what *percentage* of their program-incentivized sales they would have expected to make. For example, the primary Counterfactual question will read:

Still thinking about your sales over the past year, if Ameren Illinois’ Midstream Instant Incentives offering had not been available, what percentage of those program-incentivized <EQUIPMENT> sales would you still have expected to make? [NUMERIC OPEN END 0%-100%]

Opinion Dynamics will report on these results of this research to the Illinois NTG WG for the IL-TRM V13.0 update process with the goal of reaching consensus on a set of changes to the Protocol to address both concerns.

#### Appendix

Below, we reproduce the working version of the Midstream Free-Ridership Protocol (dated April 27, 2023) that this memo deviates from.

* 1. Midstream Free-Ridership Protocol[[1]](#footnote-1)

Typical energy efficiency programs offer incentives to end-use customers to purchase more efficient equipment. These can be referred to as “downstream” incentives, or downstream programs. Moving up the supply chain, the next entities are distributors, contractors, and design professionals. Programs aimed at influencing these market actors are referred to as “midstream” programs. “Upstream” programs target manufacturers and potentially retailers.

5.4.1 Using This Protocol

The methods described in this section should be applied for estimating NTGRs for midstream programs in which the incentives are paid directly to distributors who have the option of sharing some or all of these incentives with the end-use customers in the form of price reductions. As discussed in further detail later in this section, programs of this type influence behavior of distributors, end-users, and intermediaries such as contractors, installers, or design professionals(to various degrees).[[2]](#footnote-2)

As a result, in midstream programs where it is believed that multiple parties are aware of the utility intervention, it is desirable for evaluators to conduct research that produces multiple estimates of free ridership from different perspectives, and to combine these estimates using guidance provided in Section 5.1: Combining Participant and Trade Ally Free Ridership Scores to estimate a NTGR that is inclusive of multiple perspectives.[[3]](#footnote-3)

In these cases, care should be taken to design an evaluation approach that most appropriately reflects the theory of program influence and selects perspectives to research free-ridership from commensurate with program theory. For example, in many midstream HVAC program models, intermediaries (contractors/installers) have an influential role in the sales process with end-use customers, and to the extent that program design attempts to influence intermediary behavior as well as distributor and end-user behavior, research with these actors may be appropriate. Conversely, in other midstream program models, intermediaries may not be a target for program activity and/or may not play an influential role in the sales process, in which case research with these actors would not be appropriate.

All midstream programs should collect contact information for participating distributors and make it available to evaluators. Midstream programs may or may not collect information for other actors such as end-use customers and/or intermediaries. In cases where midstream programs do not collect customer information, end-user research will generally not be feasible, and free ridership estimates will not be able to include an end-user perspective on free-ridership. In cases where midstream programs do not collect intermediary information, intermediary research may still be feasible, but will require care in order to define an appropriate sample that accurately represents the pool of intermediaries relevant to the program. For example, some midstream programs may not track installing contractors but may maintain a list of qualified program allies who would be expected to complete a significant share of installations receiving program incentives.

If evaluation constraints do not allow for high quality research to be conducted from multiple perspectives, it is likely preferable to conduct high quality research from only one perspective rather than lower quality (e.g., minimal sample size) research from multiple perspectives, and the evaluator may choose to utilize only one approach without it being considered a divergence from the IL-NTG Methods. In this case, the evaluator should carefully consider the specific design and intent of a given program when choosing the appropriate protocol(s) for evaluation and must document the rationale for its decision in the evaluation plan.

Ultimately, the protocol(s) to be used for a given program is defined in Table 3-1 and Table 4-1. If the design of a given program changes significantly, then it may mean that the NTG protocol listed for that program in Table 3-1 or Table 4-1 is no longer appropriate. In addition, the evaluator may determine that the defined customer or distributor NTG algorithms need to be substantially modified to accommodate the specific design of a midstream program. If the evaluator chooses to use an alternative method or approach to estimate the NTG, the evaluator should follow the procedures outlined in Section 1.4: Diverging from the IL-NTG Methods. For new programs the choice of protocol(s) will be ultimately at the discretion of the evaluator.

Knowing that they will receive an incentive for selling high efficiency units, distributors may choose to increase their stock of high efficiency units, and/or to upsell high efficiency units to contractors. Distributors may also choose to offer training sessions or marketing campaigns aimed at engineers, architects, and contractors to increase awareness of these high efficiency units. As a result of the program’s actions:

Contractors/customers may be more likely to purchase high efficiency units because they are in stock,

Contractors/customers may be more likely to purchase high efficiency equipment because the distributor upsold these units,

Contractors/customers may be more likely to purchase high efficiency units because the incremental cost is lower than it would have been without the incentive (assuming the distributor uses the incentive to reduce the price of the equipment), and

Design professionals and contractors may be more likely to specify or recommend high efficiency units because they are more aware or more familiar with these options.

The expected overall outcome is that a greater proportion of customer purchases will be high efficiency units. As distributors sell more high efficiency units, manufacturers will respond by producing more high efficiency equipment. Ultimately, the overall market in a utility’s service territory will become more efficient than it otherwise would have been, or it will achieve this efficiency sooner than if no intervention had occurred.

To assess impacts from this type of program, the evaluator needs to determine if the distributor changed their practices in a way that ultimately influenced the customer’s buying decision. Assessing the influence of the program involves conducting in-depth interviews with participating distributors and asking them how they would have behaved in the absence of the program. While interviews with others such as contractors and design professionals can also be conducted in order to develop a more complete understanding of the influence of the program, the focus of the protocol in Section 5.4.2below is on the distributor interviews.

This protocol is based on the key considerations and guidelines for estimation of free ridership for non-residential programs that is described in Section 3.1.1: Core Non-Residential Free Ridership Protocol. The process to be used for scoring free ridership is described in Section 3.1.1.1: Core Non-Residential Free Ridership Scoring Algorithm. This midstream protocol can be used for estimating NTGRs for both residential and non-residential midstream programs that focus on distributors.[[4]](#footnote-4)

To ensure that the midstream NTGR approach covers all avenues of program influence, one should develop a logic model based on discussions with utility program staff, implementer staff, and a general review of midstream programs. The midstream NTGR approach recommended here is designed to be flexible as the midstream incentives may be impacting distributors’ businesses in one of many ways—including via changes in stocking, upselling, price reduction, etc. Ultimately, the midstream program should be given credit for influence via any of these causal pathways. Note that a midstream program might have longer-term impacts that are not immediately measurable. Such longer-term impacts manifest as “market effects,” which signify a transformation in the underlying structure and functioning of the market. This midstream protocol does not address the measurement of such market effects.

5.4.2 Free Ridership Estimation Methodology

This methodology uses two indicators of free ridership:

Program Influence (PI) FR Score, and

Counterfactual (CF) FR Score

These scores are then averaged to arrive at a final free ridership value. The algorithm shown in Section 3, Figure 3-1: Core Free Ridership Algorithm, can be used to calculate the free-ridership. The resulting NTGR value should be weighted proportionate to the ex post gross kWh savings for each respondent.

The one exception to the free ridership algorithm described above concerns the quantity and timing adjustment. Note that normally, in the case of downstream rebate programs, it is possible that the old equipment was still functioning, but the program induced the participant to swap out the equipment before the end of its useful life. Because of the conceptually challenging nature of a timing question for distributors, it has been removed.

This protocol starts with the Core Non-Residential Protocol methodology outlined in Section 3, Figure 3-1: Core Free Ridership Algorithm and suggests modifications to the free ridership questions to recognize the unique nature of midstream programs. Below are some examples of the types of questions that could be asked of distributors for each of the three pathways to program influence.

**5.4.2.1 Strategies Used**

First, the evaluator must ask each distributor which of the available sales strategies they used to promote program-qualified equipment.

Now, I’m going to ask you about the various strategies you might have used to sell program-qualified equipment. Please indicate which ones you have used. [READ]

\_\_\_ Upsell contractors to purchase program-qualified units

\_\_\_ Conduct training workshops for contractors

\_\_\_ Increase marketing of program-qualified units

\_\_\_ Reduce the prices of program-qualified units

\_\_\_ Increase the stocking or assortment of program-qualified units

\_\_\_ Discuss the benefits of program-qualified units with design professionals

\_\_\_ Other (Please describe: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

**5.4.2.1.1 Program Influence FR Score**

The PI Score measures the overall influence of the program on the distributor’s efforts to sell more energy efficient equipment instead of a less efficient alternative. It is developed based on a single question that asks the respondent to rate the overall importance of the program on the decision to install energy efficient equipment using a scale from 0 to 10, with 0 being not at all important and 10 being extremely important.

As described in Section 3.1, to ensure the respondent thinks about the program holistically, before asking the single question, the survey will remind the respondent of various program factors that might have influenced decision-making. This could either be through one or more questions about various program factors, or through a read-in of program factors.

Program factors are those utility actions designed to convince the distributor to increase their stock of efficient equipment and to change their sales strategies in order to sell more of these more energy efficient models. These might include such things as the incentive, information about the cost-effectiveness of the more efficient units, promotional materials, and the training of sales staff. The various program factors referenced in the survey will be determined by evaluators based on available information.

Evaluators will calculate the PI Score for each survey respondent using the following equation:

This score can range from 0 (no free ridership) to 1 (full free rider).

**5.4.2.1.2 Counterfactual FR Score**

The CF Score is based on the most likely action the distributor would have taken with respect to sales strategies for program-qualified equipment in the absence of the program.

Evaluators will administer a counterfactual likelihood survey question, which asks the respondent to identify which alternative among a list they would have been most likely to choose if the program and rebate had not been available. The scenarios will be specific to the program, but generally include options such as using fewer sales strategies to sell program-qualified equipment or using the same sales strategies as they did through the program. The respondents will be asked a follow-up question as illustrated in Figure 3-1.

In cases where the response to the initial scenario question indicates likely free-ridership, evaluators will ask a follow-up question that obtains the likelihood on a 0 to 10-point numeric scale (where 0 means “not at all likely” and 10 means “extremely likely”), absent the program, that the respondent would not have used the same sales strategies as they did through the program.

In cases where the response to the initial scenario question indicates likely non-free-ridership (flagged as “initial presumption of FR = 0” in Figure 3-1), evaluators will ask a follow-up question that obtains the likelihood on a 0 to 10-point numeric scale (where 0 means “not at all likely” and 10 means “extremely likely”), absent the program, that the respondent would have used the same sales strategies as they did through the program.

There are two options for calculating the CF Score based on the above questions.

Option 1 is the default choice and should be implemented except in cases as defined in Option 2 below. In Option 1, evaluators will calculate the CF Score for each survey respondent based on responses to the likelihood follow-up questions using one of the following equations, with the selection of the equation based on the initial scenario selected by the respondent:

Option 2 is an alternative option to be used only in cases where evaluators believe that there is evidence that respondents are not able to understand and correctly answer the likelihood follow-up questions. In Option 2, evaluators will set the CF Score equal to a presumption of FR based on the initial response to the scenario question only. Figure 3-1 provides prescribes initial presumption of FR values for default scenarios. These values are binary and take the form of either no free ridership (CF Score = 0) or total free ridership (CF Score = 1).

If evaluators choose Option 2, they should still report Option 1 scores and provide a rationale for why they believe Option 1 is not appropriate based on available data. Option 2 scores do not need to be reported if Option 1 is chosen, but evaluators may wish to report the scores to enable future dialog in the NTG Working Group as to the form of the Core Non-Residential Free Ridership Protocol.

Both options for calculating the CF Score produce scores that can range from 0 (no free ridership) to 1 (full free rider). Option 2 will lead to CF Score values that are *only* 0 (no free ridership) or 1 (full free ridership).

**5.4.2.1.3 Program Factors, Consistency Checks, and Quality Control Review**

Evaluators should also follow the guidelines regarding program and non-program factors, consistency checks, and quality control review in Section 3.1.1: Core Non-Residential Free Ridership Protocol.

Consistency checks, specifically, should be included consistent with guidance in Section 3.1.1.1.4. Additional consistency check questions to resolve identified inconsistencies remain optional, but may be used by the evaluator to resolve inconsistencies in real time if desirable.

**5.4.2.1.4 Alternative Approaches to Estimation of Midstream Influence**

The approach for assessing program impacts described in this section should not be considered exclusive or exhaustive. However, use of a different method or of a modified algorithm will be considered a deviation as discussed in Section 1.4: Diverging from the IL-NTG Methods, and will require a proposal to the Illinois SAG and approval of the proposed method by the SAG. Some additional potential methods that would be considered a deviation from this protocol will now be discussed. Within the general framework of the non-residential algorithm, there are other possible ways to construct indicators of free ridership depending on the data available. For example, for the No-Program FR Score, if the evaluator can obtain historic and current category sales data from each participating distributor, these data can be combined with program sales data (that they are required to provide to the utility) to determine the shift in efficient market shares at the distributor and program levels. If current category sales data are not available, the evaluator could ask the distributors about changes in these shares from the pre- to the post-participation periods (see example from EMI, 2018), although this approach is likely less reliable than shares based on recorded sales data. Or, one could also conduct an interrupted time-series analysis of monthly sales of program-qualified units. There may also be qualitative methodologies which can be combined with quantitative methodologies to enhance the accuracy of program impact estimates. One could also employ a theory-driven evaluation framework (Coryn, 2011) within which an evaluator could assess the program’s effectiveness, guided by the program theory and logic model. For a complex midstream distributor program, an evaluator could develop performance metrics for each activity, output, and outcome and assess the extent to which major activities of the program have been and are being successfully implemented and whether these activities had led to or are likely to lead eventually to the expected short-, mid-, and long-term outcomes. Of course, as evaluators choose to use some of these other methods, they must propose and defend a modified algorithm that can include the results from using these other methods.

1. Note that the method for assessing trade ally spillover is included in Section 5.2. [↑](#footnote-ref-1)
2. Note that the Illinois TRM does not currently prescribe a specific method for assessing free ridership from a trade ally perspective. Consistent with Section 5.1.1, until such a prescribed methodology becomes available, evaluators should coordinate on survey design and calculation algorithms for trade ally free-ridership estimation. [↑](#footnote-ref-2)
3. In cases where midstream programs require distributors to pass the entire incentive to a customer and collect customer information, it is still likely that the program is affecting distributor behavior, and distributor research is still valuable. [↑](#footnote-ref-3)
4. See Section 4.3 for a description of an approach for calculating NTG specifically for Residential Upstream Lighting programs. [↑](#footnote-ref-4)