ENERGY EVALUATION EUROPE
2020 Conference
Here East, Olympic Park, London
Monday 29 June to Wednesday 1 July 2020
Formerly IEPPEC
Accelerating the energy transition for all: Evaluation's role in effective policy making

Attributing Savings of Utility Midstream Energy Efficiency Programs: Standardizing a Protocol to Estimate Free Ridership

Laura Agapay-Read, Jan Harris, *Guidehouse, United States (laura.agapay.read@guidehouse.com)*

ABSTRACT

In North America, utility energy efficiency programs typically incentivize end use customers to purchase more efficient equipment. These are referred to as downstream programs, or rebate programs. Programs aimed at influencing market actors such as distributors, contractors, and design professionals are referred to as midstream programs. Upstream programs target manufacturers and potentially retailers.

Stakeholders access multiple benefits when a standard protocol for estimating net-to-gross (NTG) ratios for midstream programs is developed throughout a jurisdiction. Evaluators of energy efficiency programs in the State of Illinois in the US have developed such a standard protocol.[1]

Midstream programs typically influence behaviour of both distributors and end users. As a result, in midstream programs where it is believed that end use customers are aware of the utility intervention, it is desirable for evaluators to conduct research that results in both end user- and distributor-based estimates of free ridership for these programs and to combine these estimates. The combination of these perspectives should be systematic and quantitatively reflect the confidence and certainty of each perspective.

This paper describes the development, definition, and application of this standard method for estimating NTG ratios for midstream programs.

Introduction

Typical Energy Efficient Equipment Rebate Programs and Measuring Attributable Savings

For decades, some utilities in North America have run energy efficiency programs[2] that offer incentives to their customers to purchase energy efficient equipment. These programs, referred to as downstream programs or rebate programs, generally offer the incentive in the form of a rebate for the purchase of energy efficient equipment. The process generally requires the participant to purchase program-approved energy efficient equipment and to submit a rebate application with proof of purchase. The program verifies the application before sending the rebate to the participant in the form of a physical bank check, a process that may take 30-60 days. The utility then claims energy savings generally, as estimated from an engineering calculation that compares the typical annual energy consumption of the replaced inefficient equipment to that of the incentivized efficient equipment.[3] This savings is referred to as gross savings.

---

[1] The Illinois Commerce Commission (ICC) directed evaluation teams to compile and formalize standard net-to-gross (NTG) methods for use in Illinois energy efficiency evaluation (Illinois TRM v8 Attachment A page 21).

[2] Programs are generally funded by rate payers through cost recovery mechanisms, as stipulated in tariffs approved by the state utility regulator.

[3] For a description of common practices for calculating gross energy savings, see Chapter 2 of https://www.epa.gov/sites/production/files/2019-06/documents/guidebook_for_energy_efficiency_evaluation_measurement_verification.pdf

Because stakeholders of utility energy efficiency programs are often interested in the gross energy savings that a program achieves and the savings that are attributable to the program (also called additional or net savings) to assess the program's cost-effectiveness, in some jurisdictions, evaluators are charged with conducting attribution research. There are multiple methods for estimating net savings for downstream programs, these methods are detailed in literature and used in practice (Violette and Rathbun 2017). For example, in North America, a common method is to survey customers and market actors. Survey-based attribution analysis involves estimating a program's rate of free ridership (the share of savings that would have occurred absent the program) and the program's rate of spillover (savings, as a percentage of program savings, that were achieved by the program but not incentivized or tracked by the program).[4] Also referred to as net-to-gross (NTG) analysis, this research uses the following equation[5] to estimate net savings from calculated gross savings which takes into account free ridership (FR) and spillover effects (SO):

NTG = 1 – FR + SO

And because:

NTG = Net savings / Gross Savings

Then:

Net Savings = Gross Savings * (1 – FR + SO)

The survey-based method estimates free ridership[6] of downstream programs by surveying program participants (by a web, telephone, paper mail-in, or in-person survey). The survey investigates participants' decision-making of the energy efficiency purchase, often exploring two related aspects: program influence and the counterfactual (Violette and Rathbun 2017). Program influence questions attempt to measure the program's influence on the participant's decision-making on implementation of the efficiency improvement. Counterfactual questions explore the actions the participant would have taken absent the program, in particular, the likelihood that the participant would have implemented the efficiency improvement absent the program and rebate. Although there is some risk of social desirability bias and hindsight bias in self-report surveys (Ridge 2009), these can be mitigated with best practices for survey design, sampling, timing, and question wording (Keating 2009, Baumgartner 2013).

[4] See Section 3 of Violette and Rathbun 2017.

[5] A version of this equation includes market effects not already captured by spillover as an addendum. Market effects encompass "a change in the structure of a market or the behavior of participants in a market that is reflective of an increase in the adoption of energy efficiency products, services, or practices and is causally related to market intervention(s)" (Eto et al. 1996).

[6] The survey-based method is also used to estimate spillover; and while the Illinois TRM includes guidance on estimating energy efficiency program spillover, the focus of this paper is on methods for estimating free ridership.

Key Features and Importance of Midstream Programs

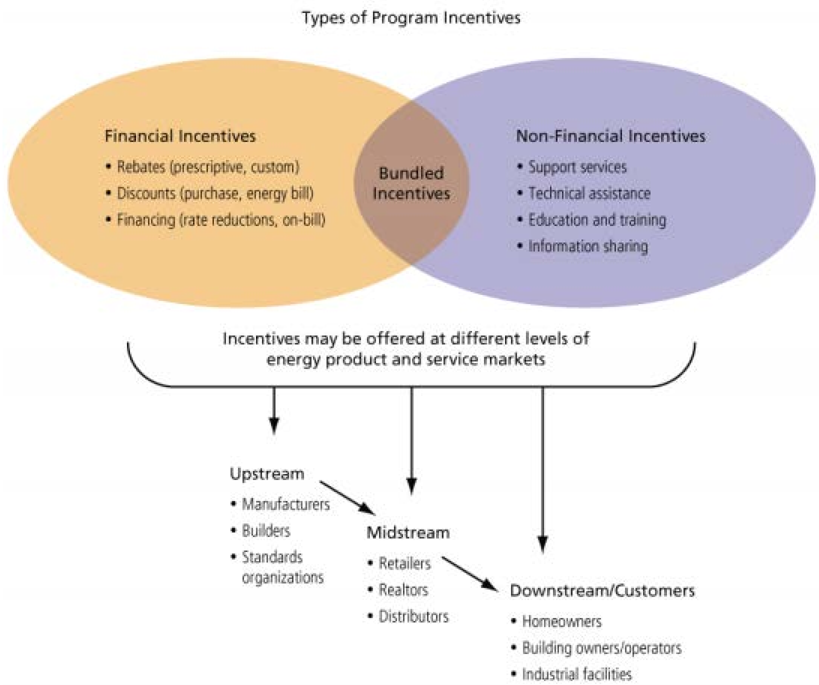

Figure 1. Program and Incentive Types. *Source:* National Action Plan for Energy Efficiency 2010.

Midstream programs[7] operate up the supply chain from end users. These programs aim to influence the behaviour of distributors, contractors, and design professionals. Midstream programs pay incentives directly to distributors that share all or some of the incentives with end use customers in the form of price reductions. The main premise of a midstream program is that the incentive for selling high efficiency units will change distributors' behaviour in one or more ways:

1. Increase their stocking of high efficiency units
2. More frequently upsell high efficiency units to contractors
3. Offer training or marketing for trade allies (engineers, architects, and contractors) to increase awareness of and promote the high efficiency units

The program theory posits that the changes in distributor behaviour will increase the availability and market share of energy efficient products to a broader contractor network. This increase in market share intensifies the likelihood that end users purchase energy efficient equipment.

This is supported by several utility midstream programs that have demonstrated substantially increased participation over their previous downstream program models (Buege, Scheidler, and Grabner 2014). Midstream programs have become more popular and more important as more utilities face ever greater energy savings goals. Table 1 lists some of the midstream programs in North America that were offered to residential[8] utility customers in 2017. The table details the programs' increase in participation over previous program years when they were structured as downstream programs.

[7] For an introduction to midstream energy efficiency programs, see https://www.michigan.gov/documents/mpsc/Navigant_Midstream_Overview_2-20-18_617578_7.pdf

[8] Midstream programs are also popular for the non-residential sector, especially for efficient lighting.

Table 1. Participation Improvement for Distributor-Focused Residential Midstream Programs Compared to Downstream Programs

Efficiency Program	Measure	Incentive Amount	Increase in Program Participation
Efficiency Maine	ENERGY STAR Certified Heat Pump Water Heater (HPWH)	$600	423% (PY1)
Efficiency Vermont	ENERGY STAR Certified HPWH	$300/$500	750%
Energize Connecticut	ENERGY STAR Certified HPWH and Natural Gas Water Heaters	$300 for gas; $600 for HPWH	1,000% (PY2)
	ENERGY STAR Certified Natural Gas Boiler and Furnaces	$450 to $800	234% (PY2)

Source: ENERGY STAR 2017.

Midstream programs benefit contractors and end use customers in that there is no program paperwork, application, or wait time for discounted high efficiency equipment because the discount and increased equipment availability are incentivized higher up the distribution channel. Although midstream programs track the number of discounted units sold, end user participant information associated with the high efficiency equipment sale is rarely tracked. To the customer, the process may appear so seamless that they are unaware of the utility program helping them. Because midstream programs aim to change distributors' behaviour in promoting energy efficient equipment, free ridership research must focus primarily on distributors. Estimating free ridership in midstream programs requires a method different from that for downstream programs.

Approaches to Estimate Free Ridership of Midstream Programs

In addition to survey-based approaches, Violette and Rathbun (2017) describe several methods suitable to estimate net savings of midstream programs. Consumption data analysis methods that use a comparison group either in a randomized control treatment design, a random encouragement design, or a quasi - experimental design can be used to directly estimate net savings. As noted in the Illinois Technical Reference Manual (Illinois TRM), consumption data analysis methods are best suited in the following situations:

- When the expected net savings per participant are large or when large participant/nonparticipant sample sizes are possible (so the difference in savings is statistically significant)
- When the program can be designed using a randomized controlled trial
- For programs where nonparticipant spillover is expected to be trivial in the comparison group
- Cases where self-selection bias[9] can be effectively controlled for

A market sales data method can also be used to directly estimate net savings. The most common approach compares post-program data with data from a non-program comparison geographical area for the same point in time, referred to as a cross-sectional comparison area method. This enables evaluators to compare the change in the program area before and after the program period as well as the change in the non-program area over the same period. The suitable application of this method requires the following considerations:

- Does an appropriate comparison geographical area exist?
- Are the market data available and complete?
- Does the program promote large numbers of homogeneous measures and have substantial influence upstream from the user?

[9] A program may attract customers who were already inclined to choose high efficiency options.

Structured expert judgment approaches are cost-effective ways to reach agreement on a NTG value when several types of evidence are available. This requires a select group of known experts that stakeholders agree have sufficient knowledge and can provide unbiased information to judge the counterfactual scenario. The process is structured in that it involves specific techniques to ensure that the experts account for key program factors, the technologies supported, and the development of other influences over time (Tetra Tech et al. 2011).

Common practice baseline assumption methods consider what typically would have been done absent the program. Evaluators determine common practices through multiple methods including self-report or on-site audits. The difference between the energy use of installed program measures and the energy use of common practice equipment is considered by some to be sufficiently close to the net savings. While there are several suitable approaches to estimating net savings of midstream programs, there is growing interest in North America to adopt standardized approaches and guidelines per jurisdiction or region.

Standardizing Estimation of Gross and Net Savings for a Jurisdiction

In the US, energy efficiency requirements for investor-owned electric and gas utilities (IOUs) vary by state. In 2019, 26 states required IOUs to achieve certain energy saving goals through energy efficiency programs (ACEEE 2019). Several states require a standard approach for calculating gross energy savings using a state TRM. Some jurisdictions have taken steps to standardize net savings research. A standard set of methodologies to estimate savings can facilitate comparing findings and assessing the attainment of regional goals that span several jurisdictions.

In 2016, the Illinois Commerce Commission directed their evaluation teams to compile and formalize consistent NTG methods for use in Illinois evaluation work. This was to ensure that energy efficiency programs across the state can be meaningfully and consistently evaluated. The Commission's directives were twofold:

1. Assess NTG methodologies and survey instruments that have been used to evaluate energy efficiency programs[10]
2. Compile the most justifiable and well-vetted methodologies in an attachment to the updated TRM (Illinois Energy Efficiency Stakeholder Advisory Group 2019)

Lead by the evaluators, the process included a (voluntary) collaboration of attribution research experts, utility program managers, program implementers, regulators, and other energy efficiency stakeholders. The working group, which met by telephone conference over several months, researched, reviewed, developed, deliberated on, and proposed standard principles and specific methods for estimating free ridership and spillover for over 12 program types, almost all of them downstream programs. Since its inception, the Illinois Stakeholder Advisory Group TRM NTG Working Group has met annually to revise and improve the standards based on evaluator experience applying them in recent research.

In anticipation of some Illinois utilities' new plans to launch midstream programs, in 2019 the Illinois Commerce Commission directed the working group to draft a standard approach to include in the TRM for estimating free ridership for midstream programs. As in 2016, this involved the voluntary contributions of a small group of program evaluators and attribution research experts (see Acknowledgments). After working independently, the subgroup proposed their method to the greater NTG working group, which reviewed, deliberated on, and revised an approach that ultimately was approved by the Illinois SAG and is presented here.

Illinois Approach to Estimate Free Ridership of Midstream Programs

The Illinois approach to estimate free ridership of midstream programs is shaped by two features: industry best practices (Violette and Rathbun 2017) that inform the Illinois free ridership methodology for downstream programs and the key drivers of program participation in midstream programs (distributors and end

[10] The original directive was to develop consistent NTG methods for every program type offered by the utilities.

use customers). The approach prioritizes collecting information on the program's influence on distributor behaviour with respect to selling energy efficient equipment. Collecting free ridership information from end user participants is of secondary importance as participant contact information is typically unavailable. When the evaluator can interview participating distributors and end users, program free ridership is calculated by combining each group's free ridership value in a way that reflects each value's likely bias, accuracy, and representativeness as best as possible.

Distributor Free Ridership

A best practice used by the Illinois TRM free ridership approach is to ask multiple questions to estimate free ridership – no single question is used to determine a respondent's free ridership (Illinois Energy Efficiency Stakeholder Advisory Group 2019, 30-31). Another best practice is to use numbers from the respondent to calculate free ridership; evaluators should not translate a qualitative response into a numeric score. The method asks the distributor(s) to numerically rate three[11] perspectives of free ridership:

1. Program Components FR Score: The extent to which (a component of) the program influenced the distributor
2. Program Influence FR Score: The extent to which all program factors influenced the distributor relative to non-program factors
3. No Program FR Score: The likelihood of the distributor making the energy efficiency changes absent the program

For each distributor, free ridership is determined by averaging the distributor's scores of the three factors. Free ridership for the program's distributors is determined by calculating a program savings weighted average for the respondents.

Figure 2 illustrates the calculation of a distributor's free ridership based on responses to survey questions for each of the three factors. It is based on the Illinois TRM's core non-residential method to estimate free ridership for downstream programs. However, there is one way that the method for midstream programs differs, it does not include a timing factor.[12] In the case of downstream 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 technical lifetime. Because of the conceptually challenging nature of a timing question for distributors, it is not included for the midstream algorithm.

[11] As noted in the Illinois TRM, different and opposing biases potentially affect two of the scores: the no program component typically indicates higher free ridership than the program component score, so combining these decreases the biases.
[12] For a description of timing factors of free ridership of non-residential downstream programs, see the Illinois TRM.

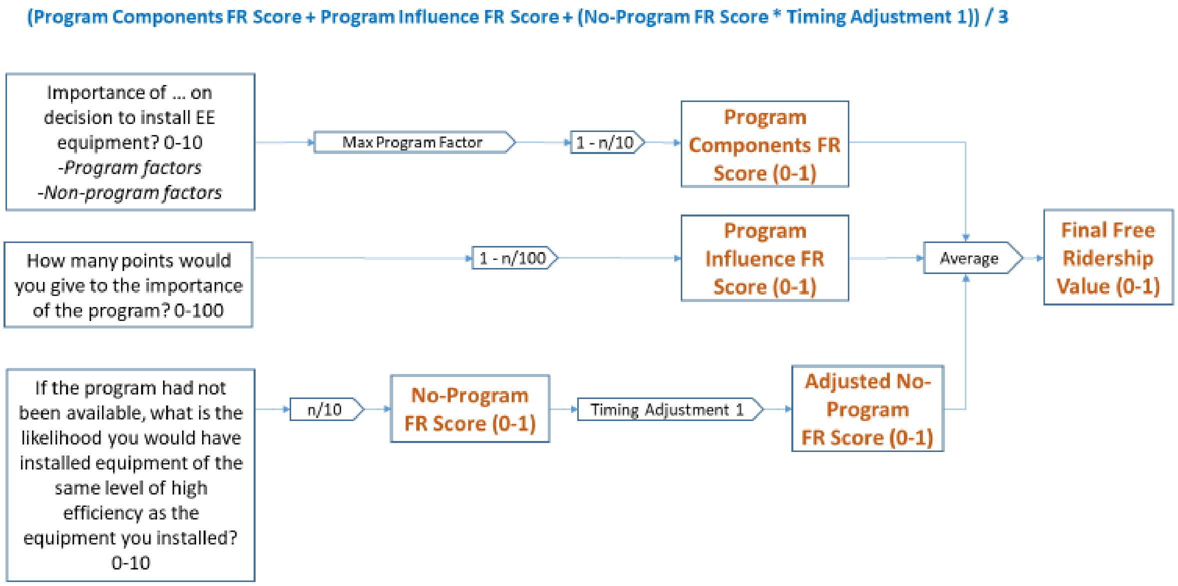

Figure 2. Illinois Core Non-Residential Free Ridership Algorithm. *Source:* Illinois Energy Efficiency Stakeholder Advisory Group 2019.

To measure a distributor's free ridership, evaluators should first identify the sales strategies used by the distributor to promote program equipment. The survey questions for the three Free Ridership Scores (FR Scores) should reference the sales strategies that the distributor employed.

Example Question:

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.

___ 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: ____)

The evaluators should next collect information for the Program Components FR Score by surveying the influence of factors on the distributor to implement the reported strategies to sell program equipment. This list of factors should include program factors and non-program factors (best practice is to use questions that rule out rival hypotheses for selling the efficient equipment) (Violette and Rathbun 2017). Program factors may include the distributor incentive, promotional materials, information on the cost-effectiveness of efficient equipment, and training sales staff. Non-program factors may include the distributor's policies to support sustainability, general concern about global warming, interest to increase sales and profits, desire to help customers reduce their energy bills, and interest in being perceived as environmentally responsible. The survey should ask the distributor to rate each factor on a scale of 0 to 10 where 0 means "not at all important" and 10 means "extremely important." The highest score reported for a program factor is used to calculate the distributor's Program Components FR Score:

Program Components FR Score = 1 – ([Maximum Program Factor Rating]/10)

The second factor, the Program Influence FR Score, attempts to quantify the importance of the program on the distributor's decision to employ the reported sales strategies for program equipment, relative to non-program factors. The evaluators ask the distributor to allocate a total of 100 points to program and non-program factors that influenced them. For clarification, evaluators should distinguish program factors from non-program factors by providing the distributor a short list of program factors and non-program factors. (Preferably, these should be taken from the respondent's top-rated factors from the Program Components question.)

Example Question:

If you were given 100 points to award in total, how many points would you give to the importance of the program factors as a group and how many points would you give to the non-program factors as a group.

The Program Influence FR Score is calculated as follows:

Program Influence FR Score = 1 – (Program Points/100)

The third factor, the No Program FR Score, explores a counterfactual scenario: the likelihood that the distributor would have used the same sales strategies to promote program equipment if the program did not exist. The evaluator asks the distributor to rate the likelihood of doing so absent the program on a scale from 0 to 10 where 0 is "Not at all Likely" and 10 is "Extremely Likely." The No Program FR Score is calculated as follows:

No Program FR Score = Rating/10

The algorithm calculates a distributor's free ridership as the average of the distributor's three free ridership Scores. The free ridership for all program distributors is calculated as the program savings weighted average of all responding distributors' free ridership rates. The Illinois TRM method also offers specific guidelines on program and non-program factors, consistency checks, and quality control review.

End User Free Ridership

The Illinois TRM midstream free ridership method acknowledges that surveying program end users may not be feasible (as the program may not collect participant end user contact information). When practical, the method suggests that evaluators collect participant end user information on free ridership. The free ridership calculation for participating end users is like the calculation for distributors. It relies on responses to multiple questions that approach free ridership from the perspectives of program influence and the counterfactual scenario. The calculation is also based on numeric scores that are provided by the respondent (i.e., the evaluator does not interpret qualitative responses into numeric scores). The Illinois TRM free ridership algorithm for residential rebate programs (with no home energy audit), shown in Figure 3, is like the one for non-residential programs; however, it does not include the Program Influence FR Score or related questions.[13] End user free ridership for a residential midstream program is calculated as the simple average of each respondent's free ridership rate, but the end user free ridership rate for a non-residential program is calculated as the program savings weighted average of each respondent's free ridership rate.

[13] This is because this question is difficult for residential customers to answer.

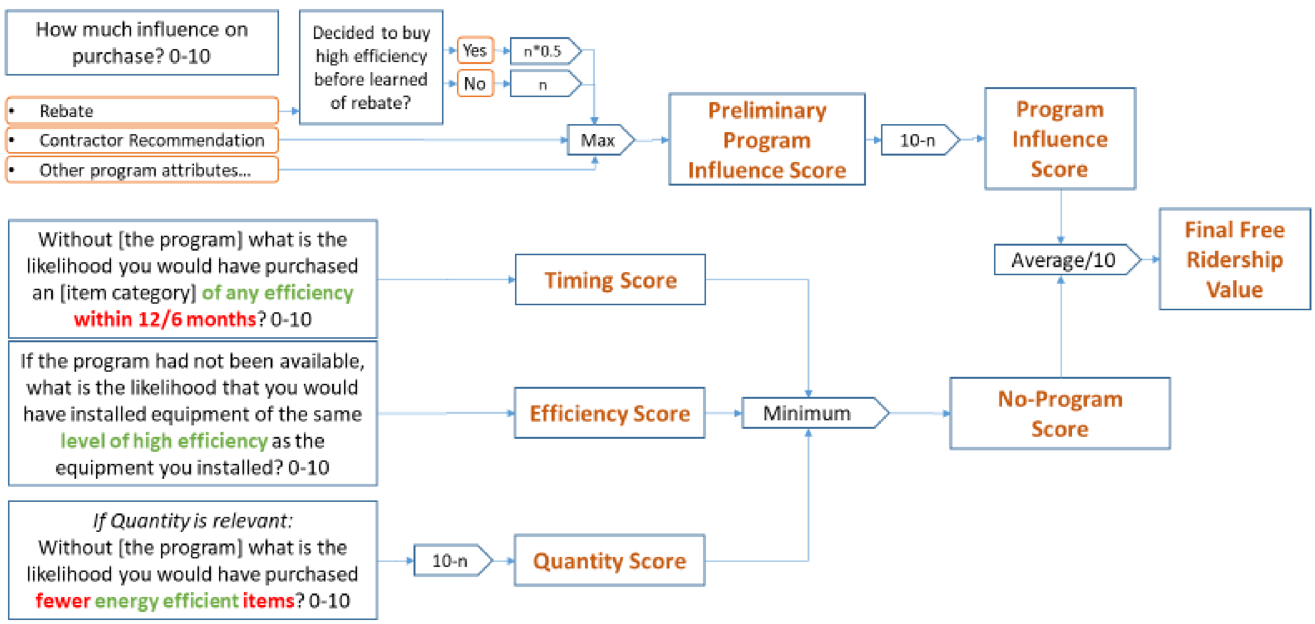

Figure 3. Illinois Residential Prescriptive Rebate (With No Audit) Free Ridership Algorithm. *Source:* Illinois Energy Efficiency Stakeholder Advisory Group 2019.

Combination of Distributor Free Ridership and End User Free Ridership

When estimating end user free ridership is feasible, the midstream program free ridership is calculated as a combination of end user free ridership and distributor free ridership. Employing a preponderance of evidence (or triangulation of results) approach that uses data from multiple sources is a best practice (Violette and Rathbun 2017). The Illinois method specifies a way for evaluators to combine these values that weights according to likely bias, accuracy, and representativeness of the results.

For example, evaluators are to combine end user and distributor free ridership results by rating the analysis methodology and data collected using their responses (rated on a scale of 0 to 10) to the following three questions:

1. All things being equal, on a scale of 0 to 10, with 0 being not at all likely and 10 being extremely likely, how likely is the approach to provide a more accurate estimate of free ridership (than the other approach)?
2. Similarly, on a scale of 0 to 10, with 0 being not at all valid and 10 being extremely valid, how valid and reliable is the data collected and the analysis performed (i.e., consider non-response bias, missing data, whether data collected was based on recollection or record keeping)
3. On a scale of 0 to 10, with 0 being not at all representative and 10 being extremely representative, how representative is the sample (accounting for sampling error [confidence and precision], and non-response bias, and any sample frame bias)?

The weight for each free ridership estimate is the average score for that estimate divided by the sum of the average scores for both estimates.

Table 2. Example Illinois Triangulation Weighting Approach

NTG Triangulation Data and Analysis	Participants	Trade Allies
How likely is this approach to provide an accurate estimate of free ridership?	6	8
How valid is the data collected/analysis?	3	5
How representative is the sample?	8	10
Average Score	5.7	9
Sum of Averages	14.7	14.7
Weight	39%	61%

Source: Illinois Energy Efficiency Stakeholder Advisory Group 2019.

Conclusion

A standard protocol for estimating free ridership in midstream programs for a jurisdiction offers multiple benefits to stakeholders. It enables consistent assessment of cost-effectiveness across utilities as well as simplifies comparisons within a jurisdiction. These benefits are magnified when the protocol is based on best practices and is designed especially for the market mechanisms of the midstream program.

Acknowledgments

The authors wish to acknowledge Jane Colby, Scott Dimetrosky, Rick Ridge, and Zachary Ross for their efforts to develop the Illinois method to estimate free ridership for midstream programs.

References

ACEEE (American Council for an Energy Efficient Economy). 2019. "The 2019 State Energy Efficiency Scorecard." https://www.aceee.org/research-report/u1908.

Baumgartner, R. 2013. "Survey Design and Implementation Cross-Cutting Protocols for Estimating Gross Savings." *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*. National Renewable Energy Laboratory.

Buege, A., L. Scheidler, and K. Grabner. 2014. "Row, Row, Row Your Commercial Lighting Program Simply Down the (Mid)-Stream?" *Proceedings of 2014 ACEEE Summer Study on Energy Efficiency in Buildings*.

Coryn, C.L., L.A. Noakes, C.D. Westine, and D.C. Schröter. 2011. "A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009," *American Journal of Evaluation*, 32(2).

EMI Consulting. 2018. *Xcel Energy: Colorado Cooling Efficiency Product: 2017 Evaluation*. Prepared for Xcel Energy.

ENERGY STAR. 2017. "Astonishing Results: Examine Surges in Participation at Distributor-Focused Midstream Programs." Distributor-Focused Midstream Programs.

Eto, J., R. Prahl, and J. Schlegal. 1996. A Scoping Study on Energy-efficiency Market Transformation by California Utility DSM Programs. Lawrence Berkeley National Laboratory. http://emp.lbl.gov/sites/all/files/lbnl-39058.pdf.

Galvin, Toben. 2018. Presentation to The Energy Waste Reduction (EWR) Collaborative. *Mid-Stream Programs: Overview*. Navigant Consulting, Inc.

Gillenwater, Michael. 2012. *What is Additionality?* GHG Institute.

State and Local Energy and Environment Program. 2019. *Guidebook for Energy Efficiency Evaluation, Measurement, and Verification*. US Environmental Protection Agency. Climate Protection Partnerships Division. Office of Atmospheric Programs. https://www.epa.gov/sites/production/files/2019-06/documents/guidebook_for_energy_efficiency_evaluation_measurement_verification.pdf.

Illinois Energy Efficiency Stakeholder Advisory Group. 2019. *2020 Illinois Statewide Technical Reference Manual for Energy Efficiency*. Version 8.0. Volume 4.

Keating, K. 2009. Paper presented at the 2009 International Energy Program Evaluation Conference. *Freeridership Borscht: Don't Salt the Soup*.

Lutz, T., P. Rathbun, and D. Violette. 2016. *Gross Savings and Net Savings: Principles and Guidance*. Northeast Energy Efficiency Partnerships.

National Action Plan for Energy Efficiency. 2010. *Customer Incentives for Energy Efficiency Through Program Offerings*. Prepared by William Prindle, ICF International, Inc. www.epa.gov/eeactionplan. Page 2.

Ridge, R., P. Willems, J. Fagan, and K. Randazzo. 2009. Paper presented at the 2009 Energy Program Evaluation Conference, Portland, OR. *The Origins of the Misunderstood and Occasionally Maligned Self-Report Approach to Estimating Net-to-Gross Ratio*.

Tetra Tech, KEMA, NMR Group, Inc. 2011. *Cross-Cutting (C&I) Free Ridership and Spillover Methodology Study Final Report.* Massachusetts Program Administrators. http://ma-eeac.org/wordpress/wp-content/uploads/Massachusetts-PAs-Cross-Cutting-CI-Free-ridership-and-Spillover-Methodology-Study.pdf.

Violette, D. and Pam Rathbun. 2017. *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures*. National Renewable Energy Laboratory.