Improving Attribution Analysis through Trade Ally Research

Jane Hummer, Navigant Consulting, Boulder, CO

Abstract

A major challenge to energy programme evaluators is discerning impacts caused by programmes (both past and present) from impacts caused by other market forces. Evaluators' "traditional" attribution analysis method of participant self-report surveys—focused on the influence of financial rebates—often fails to recognize the influences that programmes have on the structure and operations of the market itself. Participants often have a narrow, time-delineated experience of a given market, while trade allies such as manufacturers, distributors, retailers, and contractors have a much broader view of market trends and a programme's influence on their efficiency sales, including sales to customers who do not directly participate in programmes. Thus, the trade ally perspective is crucial to a thorough estimation of a programme's net impacts and progress toward market transformation.

Our team has developed a theory-based approach to collecting and analysing sales data from trade allies to estimate net impacts and increase confidence in impact evaluation results. The fundamental difference between this approach and a more traditional "net-to-gross" analysis is that the analysis starts with an estimate of total adoption of an energy efficiency measure, rather than program-reported gross impacts. This approach relies on trade allies to provide sales data that allows the analyst to estimate the total change in efficiency adoption from the preprogramme time frame to the present. The total increase in efficiency adoption then can be easily estimated. Some share of that increase (between 0% and 100%) is due to the programme's influences. The analyst's job is to narrow that range based on qualitative and quantitative evidence of the programme's influences, guided by the programme theory. Finally, the programme-influenced change in efficiency sales is compared to the number of programmerebated sales to estimate the net programme impacts.

This paper will discuss methods for developing data collection tools based on programme logic, recruiting trade allies into the study, asking the right questions to ensure unbiased and complete responses, and structuring and conducting transparent analyses. The paper will also highlight lessons learned and results from applications of this method on various evaluation projects.

The Importance of the Trade Ally Perspective

A major challenge to energy programme evaluators is discerning impacts caused by programmes (both past and present) from impacts caused by other market forces. Several different terms exist for this type of analysis, including net-to-gross (NTG), attribution, and additionality, but fundamentally they all estimate the same thing: what would have happened in the absence of the program? Note the choice of the word "estimate"; it is ultimately impossible to "measure" what would have happened in an alternate universe in which the programme did not exist. We have to try, though, because attribution is an important consideration when deciding how to spend limited programme funds. No one wants to invest in a market that would have achieved the same outcomes even without the programme investment. Luckily, with a

carefully structured analysis that captures a broad range of perspectives on the market, we can increase our confidence in our estimates and guide stakeholders toward the most effective use of programme dollars.

This paper focuses on the use of sales data and qualitative insights from trade allies (such as manufacturers, distributors, retailers, and contractors) to improve attribution analysis. The most common approach to attribution analysis in the United States is the self-report survey of participating end-use customers. These surveys focus primarily on the influence of financial rebates and often fail to recognize the influences that programmes have on the structure and operations of the market itself.¹ Participant surveys are also subject to a number of psychological biases that some evaluators argue result in an overestimate of free ridership.²

The failure of participant self-report surveys to capture the full range of programme influences is particularly troubling in programmes with midstream or upstream elements. Midstream or upstream programme strategies target trade allies such as contractors, distributors, retailers, and manufacturers with financial incentives, education, marketing assistance, and/or other interventions. End-use customers are often unaware of these programme strategies, and thus cannot consider them when answering free ridership questions in a survey. The authors of the Uniform Methods Project's chapter on *Estimating Net Savings* (Violette & Rathbun, 2014) went so far as to describe the use of participant self-report surveys on midstream and upstream programmes as "inappropriate."

In short, participants typically have a narrow, time-delineated experience of a given market, while trade allies have a much broader, multi-year view of market trends and a programme's influence on their efficiency sales, including sales to customers who do not directly participate in programmes. Thus, the trade ally perspective is crucial to a thorough estimations of a programme's net impacts and progress toward market transformation.

Fundamentals of Net-to-Gross

Net-to-gross is the ratio of savings the programme *influenced* to savings the programme *paid for*. Programme records provide the number of program-rebated high efficiency sales for each trade ally, but trade allies often have additional sales of high efficiency sales that occur

¹ Tom Eckman argues that our industry's focus on free ridership ignores the many years (in some regions, decades) of utility investment in creating the infrastructure to market and deliver energy efficiency products and services to end-use customers. He describes this folly as such: "In this scenario utilities built the road, bought the bus, hired the driver, paid for the gas, and then paid someone with no means of transport to take a ride to visit her dying grandmother. A month later, an evaluator asks this person: 'Would you have taken that ride to visit your dying grandmother if you had to pay full fare?' If she says 'yes,' the investments in the transit system are no longer cost-effective because she's a free rider. This assumption is at best overly conservative and more likely just plain illogical." (Eckman, 2011)

² Several psychological theories predict that participants answer questions in self-report surveys in a manner that will result in overestimates of free ridership (Peters and McRae, 2008). Attribution theory says that the very act of making a decision changes how one views themselves, indicating that it is very hard for people to imagine themselves making any other purchase decision than the one they already made. Cognitive dissonance theory says that people feel most comfortable when their actions are consistent with their stated beliefs, and being asked questions about why they took a specific action (e.g., purchased an energy-efficient appliance) that they had not intended to purchase may trigger a cognitive dissonance that they resolve by claiming that they had always intended to choose the energy-efficient option. Finally, participants may view making an energy-efficient purchase as a socially desirable behavior and have a motivation to present themselves as an environmentally conscious consumer in their response to survey questions.

outside of programmes. All sales of high efficiency equipment fall into four categories based on whether the sales were program-influenced or not, and whether they were program-rebated or not. Free ridership (rebated) and naturally occurring baseline activity (non-rebated) are not included in estimates of net impacts; program-influenced participants and program-influenced non-participants (spillover) are included in the net impacts.

Figure 1 summarizes these four categories and their relationship to the net and gross savings estimates.





Market-Focused Attribution Analysis

A major difference between our approach and a more traditional net-to-gross analysis is that our analysis starts with an estimate of the participating trade allies' total sales of an energy efficiency measure, rather than only the program-reported gross impacts. Our approach collects trade ally sales data that allows the analyst to estimate the total increase in efficiency adoption from the pre-programme time frame to the present. Some share of that increase (between 0% and 100%) is due to the programme's influences. The analyst's job is to narrow that range based on qualitative and quantitative evidence of the programme's influences, guided by the programme theory. Finally, the program-influenced change in efficiency sales is compared to the number of program-rebated sales to estimate the net programme impacts.

Change in Unincented Sales over Time

Trade allies that participate in programmes are usually able and willing (with a guarantee of confidentiality and sometimes a small financial incentive) to share details on their sales of high efficiency measures over time. Trade allies are asked to provide data on all high efficiency

sales, not only those that receive programme rebates, for the last year *before* they participated in the programme as well as the programme year(s).

The change (typically an increase) in high efficiency sales over time is easily measured. Most trade allies report some level of high efficiency sales prior to their involvement in the program; our analysis assumes that this level of sales forms the *minimum* baseline for the market, meaning we would expect that in subsequent years, the trade ally would continue selling *at least* that volume of high efficiency units even without the program.³ Our analysis focuses on the increase in high efficiency sales, and the percentage of that increase that was influenced by the program, as shown in **Figure 2**.



Figure 2. Change in High Efficiency Sales over Time

A comparison of the trade ally's sales data to the programme records provides a preview of whether the trade ally's NTG is likely less than 1.0 (indicating free ridership) or potentially more than 1.0 (indicating possible spillover), as shown in **Figure 3**. If unincented sales (i.e., high efficiency sales that did not receive rebates) are *less* than the minimum baseline established by the pre-programme sales level, there is likely free ridership. If the unincented sales exceed the minimum baseline, there is possible spillover, pending further exploration of the baseline. The likelihood of spillover is less certain, because the baseline may not be frozen at the pre-programme levels, due to other market forces driving increased efficiency.

³ The assumption of a minimum baseline based on pre-programme sales level may be flawed in certain circumstances such as during an economic recession. Qualitative questions in the trade ally interviews can explore those potential circumstances.



Figure 3. Comparison of Unincented Sales to Minimum Baseline

With this data, the analyst can determine the minimum estimate of free ridership if unincented sales in the programme year are less than the pre-programme sales of the high efficiency measure. This is based on that assumption that—in the absence of the program—high efficiency sales would have continued at the pre-programme levels at a minimum. **Equation 1** demonstrates the calculation of the minimum free ridership estimate; note that if unincented sales in the programme year exceed pre-programme sales, the minimum free ridership is zero.

Equation 1. Minimum Free Ridership Equation

$$Minimum FR = Maximum \left(\frac{Pre Program Sales - Unincented Sales}{Incented Sales}, 0\right)$$

However, it is possible that even in the absence of the programme, more customers would have been purchasing high efficiency units than during the pre-programme time period due to other naturally occurring market forces (e.g., changing codes and standards, increasing energy prices, decreasing prices of high efficiency options, etc.). The analyst must develop an attribution factor (between 0% and 100%) to estimate what share of the increase in high efficiency is due to the programme and not these other market forces. The next section of this paper will discuss options for estimating that attribution factor, but for now just assume that analysts can develop an attribution factor. The attribution factor can be used to estimate the maximum free ridership (**Equation 2**) as well as spillover (**Equation 3**). Note, in some analyses we develop separate attribution factors for free ridership and spillover, but in others we use the same attribution factor to calculate both free ridership and spillover, depending on the programme logic and data collection constraints.

Equation 2. Maximum Free Ridership Equation

$$Max FR = Minimum FR + \frac{Increase in Total Sales * (1 - Attribution Factor)}{Incented Sales}$$

Note if the attribution factor is 100% (meaning the programme is solely responsible for the increase in high efficiency sales since the pre-programme baseline), the minimum and maximum free ridership estimates are identical.

Spillover exists only when the trade ally reports that they are selling more high efficiency units without programme incentives than they sold in the pre-programme time period, and when there is sufficient evidence to attribute those additional sales to the program. Thus, spillover is equal to the *increase* in the sale of unincented units, multiplied by an attribution factor, and then divided by the number of incented sales to estimate the spillover as a percentage of program-incented units (**Equation 3**).

Equation 3. Spillover Equation

Spillover = <u>Increase in Unincented Sales * Attribution Factor</u> <u>Incented Sales</u>

Attribution Factor

The attribution factor is an estimate of the share of the increase in energy efficiency sales that would *not* have occurred in the absence of the programme. It is unknowable. Measurement—as opposed to estimation—of the attribution factor would require us to build a time machine, go back in time, and measure the high efficiency sales that would have occurred in an alternative universe that is exactly the same as ours in every way except that the programme did not exist. That said, there are ways to make a defensible estimate of an attribution factor for the purposes of these analyses.

The first step is to ensure that the evaluation team has a thorough understanding of the programme theory and logic⁴, which may require interviews with programme staff and review of programme documents. An effective programme theory and logic answers the following questions:

- Who is the programme attempting to influence? (End-use customers, retailers, contractors, distributors, manufacturers, etc.)
- What are the programme's intended outcomes? What market barriers are they trying to overcome?
- What types of strategies and interventions are employed to achieve these outcomes?
- What are the indicators of programme success?

The goal is to identify (and ask questions about) all of the ways in which the programme may have influenced the trade allies and removed barriers to the adoption of high efficiency measures. Eto et al. (1996) identify a range of possible programme influences on trade allies, including:

- Changes in promotional practices
- Changes in business strategies

⁴ A logic model is a powerful visual tool for anyone who is designing, implementing, or evaluating a programme. Logic models help ensure that all parties share a common understanding of the programme's strategies and intended effects on the market (W.K. Kellogg Foundation, 2004).

- Changes in prices offered to customers
- Creation of new players
- Changes in stocking and distribution practices
- Changes in design practices
- Changes in service offerings
- Development of new skills

These influences may result from programme strategies such as technical training, customizable marketing materials, marketing assistance, inclusion in a trade ally directory (leading to more project leads), direct referrals of project leads, etc. For each of these strategies, the interview guide should first ask whether or not the trade ally benefitted from any of these programme strategies. If they say yes, then the interviewer asks an open-ended question about *how* they benefitted, followed by an influence question in which the trade ally rates the influence of the programme's strategies on their high efficiency sales on a 1-5 or 0-10 scale. The response to the influence question is then translated into a 0-100% influence score.

Direct questions about free ridership or attribution should be asked *after* the programme influence questions, so that trade allies are first primed to recall all of the ways in which the programme may have aided their business. A direct question about attribution might be framed as "In your opinion, what percentage of your total [efficient measure] sales in [programme year] would have occurred even if you had not participated in the programme?" The response is inversely related to attribution; a high percentage indicates low attribution/programme influence and a low percentage indicates high attribution. An alternative question could be "How likely is it that you would have sold the same volume of this product if the programme did not exist?"

The interview guide should build in a consistency check procedure in which trade allies who indicate both low attribution *and* high programme influence (or vice versa) are asked to clarify their responses. This ensures that trade allies understood the questions properly and provides an opportunity to collect nuanced insights on attribution through open-ended questions. There may be programme influences experienced by the trade allies which the programme logic did not predict or other market influences that were more significant than the programme.

Our evaluation team has used different algorithms to estimate the final attribution factor depending on the programme type and the available data. The most important thing is to be transparent in the formation of the attribution factor and to consistently apply the same algorithm to all trade allies. In some cases, the attribution factor used in the NTG analysis is based on the trade ally's direct estimation of the attribution factor and the programme influence questions are simply used as a consistency check. In other cases, the average or maximum of the programme influence scores may form the attribution factor. The results of the process evaluation (e.g., the assessment of the effectiveness of programme marketing materials or trade ally training sessions) may aid in designing the attribution factor algorithm.

Implementation of the Approach

Our evaluation team has implemented variations on this approach for a range of energy efficiency programmes, including two residential HVAC programmes, two commercial lighting programmes, a commercial HVAC programme, and a commercial gas programme. This section will discuss the practical implementation of this approach as well as some lessons learned.

Data Collection

As discussed previously, before any data is collected from trade allies, the evaluation team needs to conduct a careful review of the programme theory and logic model to identify the programme's strategies and expected outcomes. The interview guide development should be a collaboration between the process evaluators and the impact evaluators, to ensure that the questions cover the full range of programme influences as well as the technical details necessary to develop a robust net-to-gross estimate.

The previous sections discussed the key questions to include in the trade ally interview guide regarding programme influences, pre-programme sales, and total sales in the current programme year. A few additional considerations for interview guide design and implementation include:

- Structure the interview guide logic to skip over questions regarding measure categories which comprise a small share of the trade ally's total sales; perhaps focus the interview guide on the 3-4 measures that the trade ally sells most often.
- Focus questions on specific measures rather than broad measure categories.
- If time allows, ask about sales of standard efficiency equipment within the same measure categories (e.g., ask about T12 lighting in addition to T8, T5, and linear LED lighting).
- Include questions about *why* high efficiency sales don't go through the programme.
- Include questions about what share of high efficiency sales go to customers outside of the utility territory; most utilities cannot count those projects as spillover.
- Design the sample to target the largest trade allies (i.e., the trade allies that comprise the majority of programme savings).
- Assure the trade allies that their sales data will not be reported individually, but analysed in aggregate with other trade allies' data.

If the programme includes a wide variety of measures, it may be more efficient to ask trade allies to complete a spreadsheet data collection form that asks about sales volumes and programme influences at the measure level. **Figure 4** presents an abbreviated version of such a data collection form used on a commercial lighting programme evaluation. In this evaluation, we offered the trade allies a \$100 incentive to complete the data collection form in addition to the telephone interview, due to the time required of the trade ally. These financial incentives typically increase the response rate of the interview effort.

Lightin	z Technology	Have you sold this specific technology to C&I customers within the past 5 years? yes/no	Before participating in the DTE program, did you recommend this technology to your customers? yes/no	Approximate Volume of Sales in 2009 (pre- program) # of lamps/ fixtures/ controls	Approximate Volume of Sales in Past 12 Months #oflamps/ fixtures/ controls	How likely is it that you would sell the same volume of this technology if the program did not exist? Rate on 1-5 scale, 5 = very likely	% of Sales that did not Receive Utility Incentives (in Past 12 Months) % of lamps/ fixtures/ controls
	T12 lamps						
	Standard T5 lamps						
Linear Fluorescent (not high bay)	High Output T5 lamps						
	Standard, or "700 Series" T8 lamps, 32W <2950 initial lumens						
	High Performance, or "800 Series" T8 lamps, 32W ≥ 3,100 initial						
	lumens						
	Low Wattage (LW) 4-foot Linear Fluorescent lamps (<32W)						
	T12 or T8 4ft Lamp Removal (in combination with HPT8 or LWT8 ballast retrofit)						
	Other (Specify type in Notes field)						
Interior High Bay	Standard T5 (High Bay Fixtures)						
	High Output T5 (High Bay Fixtures)						
	High Performance, or "800 Series" T8, 32W ≥ 3,100 initial lumens						
	(nigii bay rixtures) High Output TS (High Pay Eixtures)						
	IFD (High Bay Fixtures)						
	42W 8 Lamp High Bay Compact Fluorescent Fixture						

Figure 4. Example of a Detailed Sales Data Collection Tool

Analysis and Results

The complicated part of this approach is asking the right questions of the trade allies. The analysis itself is transparent and simple. Each measure is analysed separately. For each trade ally/measure combination, you need four pieces of information:

- Pre-programme sales of high efficiency measure (unit count)
- Total sales of high efficiency measure in programme year (unit count)
- Programme-incented sales in programme year (unit count)
- Attribution factor

Using the equations presented above in the Market-Focused Attribution Analysis section, free ridership, spillover, and net-to-gross (NTG) ratios are estimated for each trade ally/measure combination. The trade ally/measure level NTG estimates are rolled up to the programme level using savings-weighted averages, as shown in **Figure 5**.



Figure 5. Estimation of Programme Net-to-Gross

Table 1 presents an example of this approach applied to a commercial and industrial energy efficiency programme. In this evaluation, fifteen participating lighting contractor trade allies completed phone interviews and a detailed spreadsheet providing unit sales data for a wide variety of lighting measures. The analysis found significant spillover (an increase in unincented sales attributable to the programme) from five measures, with savings equivalent to 29% of total programme savings. The interviewed trade allies linked their increased sales of LEDs and low wattage linear fluorescents outside of the programme to the programme's influences on their knowledge of the technologies and changes in their stocking practices.

	Change in Unincented Sales	Per-Unit Savings (kWh)	Total Savings (kWh)	Attributable Savings (kWh)
Measure	Α	В	$C = A^*B$	D = C* Attribution Factor
LED Recessed Down Lights	30,635	195	5,973,825	2,365,691
Exterior LEDs	26,400	268	7,075,200	2,606,300
PAR LED	57,815	116	6,706,540	1,932,183
LED A-Line	18,777	196	3,680,292	1,129,715
Low Wattage Linear Fluorescent	110,500	29	3,204,500	978,750
Total Trade Ally Spi	10,565,494			
Total Programme Sav	22,977,159			
Trade Ally Spillover	46%			
Lighting Savings as %	63%			
Trade Ally Spillover	29%			

Table 1. Example of Spillover Estimation Results (Commercial Lighting Programme)

This approach generates a rich set of qualitative data in addition to the NTG results, which should be included in the reporting. Supplementing the quantitative analysis with direct quotes from the trade allies on the programme's influences provides valuable context for the NTG results.

Limitations and Lessons Learned

This approach can be scaled and modified for a number of different programme types, but it works best when the majority of programme savings come from a small number of measure types with low variability in measure-level energy savings estimates. The approach can be unwieldly for programmes with a heterogeneous mix of measure types, particularly custom measures for large commercial and industrial customers, due to the sheer volume of data that would need to be collected, as well as the uncertainty in the energy savings of unincented measures relative to programme-incented measures. Even in programmes with low variability of energy savings within the programme, it is important to consider whether unincented measures have the same baseline (and thus the same savings estimates) as programme measures. Future iterations of our data collection tool will include questions regarding the magnitude of spillover savings relative to programme projects.

One challenge we encountered was related to programme data tracking. Full implementation of this approach requires accurate unit counts at the measure level for each trade ally's programme-incented projects. The implementer of one programme was only able to provide broad measure category descriptions of projects (e.g., "lighting" or "boiler"), rather than detailed descriptions with measure counts such as "250 13W compact fluorescent light bulbs." For this programme, the evaluation team had to base the interview questions on high level measure categories and project counts rather than measure counts, which both limited the value of the qualitative data on market trends and resulted in an unacceptable level of uncertainty regarding the energy savings of spillover projects. This experience was a good reminder to evaluators to work with utility programme staff and implementers early on to ensure that programme data will be collected and stored in a manner that enables an effective evaluation.

A significant shortcoming of this approach *as implemented to date* is that we have limited our data collection to participating trade allies.⁵ Thus, the analysis has assumed that the programme has not had any influence on non-participating trade allies' sales of high efficiency equipment, and this assumption may be incorrect. Programmes may have both positive and negative effects on non-participating trade allies' sales of high efficiency equipment. For example, a programme with a broad advertising campaign may increase consumer awareness of and interest in high efficiency options, and that may result in non-participating trade allies selling more high efficiency equipment, which could be considered spillover. However, it is also possible that customers who would have purchased high efficiency on their own (outside of programme influence) are being drawn away from non-participating trade allies' pre-programme high efficiency sales are a minimum baseline is flawed and free ridership is higher than our analysis would indicate.

Expanding the data collection to include any significant non-participating trade allies in the utility's service territory would reduce uncertainty in the NTG analysis. This would also provide a valuable opportunity for process evaluators to explore why certain trade allies choose not to participate in the programme. Note, however, that non-participating trade allies may require a more substantial financial incentive to participate in the study than participating trade allies typically require, as non-participating trade allies may have less motivation to remain in the programme's good graces. Another valuable input into this analysis (particularly for residential programmes) would be periodic general population surveys to assess changes in consumer awareness of and interest in the utility's programmes and energy-efficient technologies.

⁵ Note that Navigant has conducted similar research with non-participating trade allies in Momentum Savings analyses for the Bonneville Power Administration. The Momentum Savings analyses are conceptually similar to this approach, but do not attempt to estimate an attribution factor, instead focusing on quantifying the total increase in energy efficiency adoption in the Pacific Northwest region of the United States (Bonneville Power Administration, 2015).

Conclusions

Trade ally research provides valuable insights into programme performance and greatly improves our understanding of net savings. This approach helps to counteract the widely recognized biases of participant self-report surveys, and it is consistent with net-to-gross estimation best practices as laid out in the Uniform Methods Protocol (Violette and Rathbun, 2014):

- The approach considers a wide range of programme influences on both customers and trade allies, not only rebates.
- Data collection and analysis occurs at the measure level, rather than project level.
- The interview guide employs a combination of quantitative and qualitative questions, as well as consistency checks to ensure that respondents understand the questions and are providing robust data.
- The analysis results in ranges, rather than point estimates, of net-to-gross ratios, and the results are savings-weighted.

Many programme evaluations already include interviews with participating trade allies as part of their normal process evaluation activities, and many of these interviews already include qualitative assessments of the programme's influences. This approach can greatly enhance the value of the information obtained through those interviews, simply by carefully structuring questions around the programme logic to feed into the net-to-gross analysis.

References

- Bonneville Power Administration. *Overview of Momentum Savings*. Prepared by Navigant Consulting, Inc. and Cadeo Group. October 2015. <u>https://www.bpa.gov/EE/Utility/research-archive/Documents/Momentum-Savings-Resources/Overview on Momentum Savings white paper.pdf</u>.
- Eckman, Tom. 2011. "Some Thoughts on Treating Energy Efficiency as a Resource." *ElectricityPolicy.com*. <u>http://www.electricitypolicy.com/Eckman-5-12-11-final.pdf</u>
- Eto, Joseph, Prahl, Ralph, and Jeff Schlegel. 1996. *A Scoping Study on Energy-Efficient Market Transformation by California Utility DSM Programs*. Prepared for the California Demand-Side Measurement Advisory Committee. <u>https://eaei.lbl.gov/sites/all/files/lbnl</u>-<u>39058.pdf</u>
- Violette, Daniel M. and Pamela Rathbun. 2014. "Chapter 23: Estimating Net Savings: Common Practices." *The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Methods*. Golden, CO: National Renewable Energy Laboratory. <u>http://www.nrel.gov/docs/fy14osti/62678.pdf</u>
- W.K. Kellogg Foundation. 2004. Using Logic Models to Bring Together Planning, Evaluation, and Action: Logic Model Development Guide. <u>https://ag.purdue.edu/extension/pdehs/Documents/Pub3669.pdf</u>