Customer Truth Serum: Comparing Customer Responses from Rebate Applications and Customer Surveys

Katherine Johnson, Johnson Consulting Group, Frederick, MD
Shaylyn Dean, Spire, St. Louis and Kansas City, MO

ABSTRACT

Determining the accuracy of customer survey responses is one of the most confounding areas of attribution research. Impact evaluations rely heavily on customers accurately recalling the influence that rebates and contractors had on making energy efficiency purchases. But are customers actually recalling these experiences correctly, and is there any way to measure these responses over time?

Two recently completed impact evaluations examined this question by comparing customer responses at two critical points: at the initial rebate application and during the customer survey conducted six months to a year after installation. This analysis was performed for two separate program cycles. This multi-year approach identified that while most customers provide consistent responses, a group of program participants switched their responses from the initial rebate application to the customer survey. So as the old courtroom adage goes, *"Were you lying then, or are you lying now?"*

This paper compares the initial rebate application answers to the same question asked during the customer survey. It also identifies the number of customers who actually "flipped" or gave inconsistent responses. Comparing the initial rebate application results to the follow-up customer survey results during both program periods identified several emerging trends:

- Most customers provided consistent responses to both question sets, suggesting that customers answer these questions truthfully.
- But there is a small cluster of customers who provided contradictory answers.

This comparison illustrates the importance of gathering customer feedback multiple times during the evaluation cycle to determine the overall results accurately.

Introduction

Determining the accuracy of customer survey responses is one of the most confounding areas of attribution research. Impact evaluations rely heavily on customers accurately recalling the influence that rebates and contractors had on making energy efficiency purchases. But are customers actually recalling these experiences correctly, and is there any way to measure these responses over time? This paper illustrates the challenges associated with accurately documenting program participant responses over time. It compares participant responses to free ridership questions during two points: the initial rebate application and answers to follow-up questions conducted months after the product installation.

Spire Inc. is a natural gas company that serves more than 1.7 million customers in multiple states. The results cited in this paper are from two independent program evaluations conducted for Spire's operating companies in Missouri: MOE (Missouri East aka Laclede) and MOW (Missouri West aka Missouri Gas Energy). Spire program staff hired the Johnson Consulting Group team to complete a process and impact evaluation of its Residential Water Heating and Space Heating Program in 2017 and 2020.

Methodology

Initially, the program evaluation team planned to assess ridership using the Net-to-Gross algorithm in multiple jurisdictions. This is a complicated set of questions that ask participants to answer multiple questions regarding the factors that influenced their purchase decision. The following figure illustrates the question logic used for this question set.

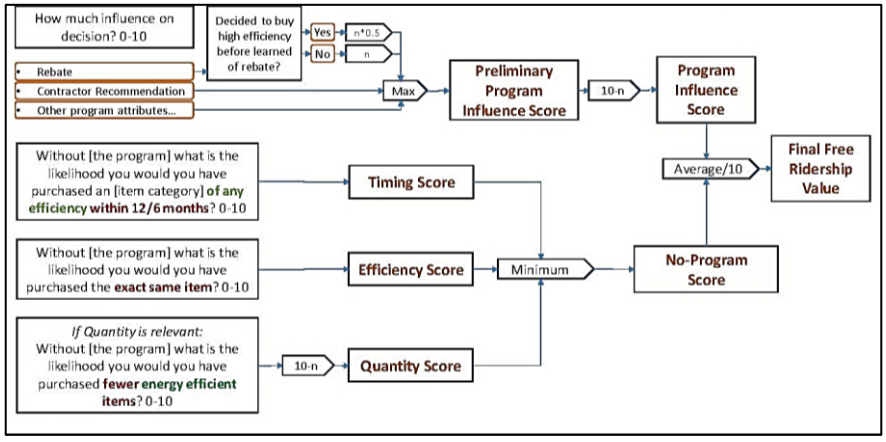

Figure 1: Free Ridership calculation methodology. *Source:* Illinois TRM Version 6, Volume 4, Figure 4-6, p. 70

The evaluation team used two approaches to measure free ridership to determine the Net-to-Gross Ratio (NTGR) for its residential program. The first approach used a standard set of NTG survey questions that are widely used for impact evaluations. However, this approach had two significant shortcomings:

- First, it appeared to overstate program free-ridership through a complicated algorithm that weighted different components which may influence participant purchase decisions;
- It did not have a mechanism to include additional free ridership data captured in the rebate application from program participants.

The overall goal of any program evaluation is to gather data from multiple sources; however, the NTG algorithm draws exclusively from participant survey data collected several months after the measure has been purchased and installed. Industry best practices have long recognized the need to move to a "fast feedback" survey, in which NTG data are collected shortly after program participation.

But Spire program staff developed an approach that is gathered NTG data at the time of the participant's actual decision-making through its rebate application. The team added two questions designed to assess free ridership in its standard rebate application. The application also includes other questions, such as the source of awareness, to provide real-time feedback regarding the effectiveness of various marketing and outreach tactics.

Initially, the evaluation team used the NTG algorithm to conduct free ridership analysis. However, the overall response rates to this question battery were low and declined further as the participant answered subsequent questions.

However, the customer survey calculated free ridership rates were also high, thus reducing program overall savings. The program weighted average for free ridership was 49 percent, with spillover of 11 percent. However, this free ridership is higher than the rate reported by customers at the time of the program application. A companion free ridership analysis compared the participant survey responses with the original answers regarding the influence of the program rebate found that free ridership varied from 23 to 44 percent. (Johnson 2017, p. 2)

Given the high level of uncertainty associated with this response rate for the free ridership questions, the program manager provided the evaluation team with the complete set of rebate applications for each evaluation period. This program database included capturing the initial responses to the simple question from the rebate application: *Did the rebate influence your decision?*

Table 1 summarizes the response rate for this question for both evaluation time periods. As this table shows, the number of responses each year was markedly higher than the survey participants (n=71, 2017; n=73, 2020, respectively).

Table 1: Summary of Responses for Rebate Application Questions from Program Participants

	2017 Results			
Rebate Influenced Decision?	Laclede	MGE	Total	Response Rate
Yes	3,897	2,945	6,842	40%
No	1,665	1,640	3,305	19%
Blank	5,573	1,238	6,811	40%
Total	11,135	5,823	16,958	100%
	2020 Results			
Yes	1,284	4,227	5,511	23%
No	561	2,462	2,923	12%
Blank	15,391	477	15,868	65%
Total	17,236	7,136	23,372	100%

As Table 1 shows, a significant number of program participants initially left this question blank, which suggests that these respondents did not feel "pressured" to answer this question. Due to a turnover in program database providers, not all rebate application data were captured in 2020. However, the sheer number of responses did provide a clear indication that free ridership rates were relatively low, based on just these responses (i.e., 19% and 12% respectively, of those participants who said that the rebate did not influence their purchase decision.)

Comparison of Responses

As a way to reconcile the discrepancies between the relatively high free ridership rates from the survey findings (49% in 2017) and the relatively low free ridership rate from the customer rebate applications (19% in 2017), the evaluation team compared the participants' initial rebate responses to their answers on the follow-up customer survey. Specifically, the team examined the participants' responses to the rebate influence question on the program application to the results from three of the questions from the follow-up customer survey:

Commented [KJ1]: Based on the customer responses from both the participant survey and application, net-to-gross for the Residential Program is calculated as follows:
- Free Ridership= 28%
- Spillover= 12%

- *FR2. How influential was the availability of the rebate from* [UTILITY] on your decision to install the **[MEASURE]** that you installed? Please use a scale from 0 to 10 where 0 means "not at all influential" and ten means "very influential."]
- FR8. *Without the [PROGRAM NAME] program rebate, how likely is it that you would have purchased the exact same* <MEASURE>? Please use a scale from 0 to 10, where 0 means "not at all likely" and ten means "very likely."

Caveats: It is important to note a few caveats with this analysis. First, these findings are qualitative due to small response rates for the survey free ridership questions, and the purpose of this analysis is to identify *actual free riders*. Furthermore, this analysis was conducted to *identify data trends* and to determine if the free ridership questions are being answered consistently from the rebate application to the follow-up survey questions. This approach focused on analyzing responses from who were either "Low Influence" ("1-3" responses) or "High Influence" (i.e., responses "8-10") for each time period.

Results

Table 2 provides a cross-tabulation of the responses to two separate questions:

- The rebate application question: "Did the rebate influence your decision?" with
- The survey question assessed the "level of influence" the rebate had on a scale of 0 to 10.

Clear patterns emerged when comparing these results over time. First, there was a remarkable amount of consistency among participants who initially indicated they were free riders (i.e., "No") respondents who also reported that the rebate had "little to no influence" on their purchase decision '(i.e., scores "0-3").

But several discrepancies emerged, which are bolded and italicized n the following table. For example, three respondents in 2017 and eight in 2020 provided contradictory findings. First, they indicated that their rebate did influence their purchase decision on the rebate application but subsequently rated the level of influence of the rebate on their decision a "0". These 11 respondents essentially flipped their responses over time, which of course, affected the overall free ridership rating.

Table 3: Illustration of Contradictory Findings Regarding Rebate Influence Over Time

Source		Influence of Rebate (Score = 10)	
Spire Rebate Application	Did Rebate Influence Purchase Decision?	2017 (n=28)	2020 (n=33)
	Yes	7	1
	No	***0***	***0***
	Blank	4	0
Spire Rebate Application		Influence of Rebate (Score =0)	
	Yes	***3***	***8***
	No	6	5
	Blank	6	3
Contradictory Findings		3	8
% of Contradictory Responses		*11%*	*24%*

The evaluation team found a similar level of discrepancies of respondents ratings over time in the analysis of the following questions, which is summarized in the following cross-tabulation:

- The rebate application question: "Did the rebate influence your decision?"
- The likelihood of purchasing the exact same measure without the rebate

This question assumes that program free-riders provide a higher rating for this question than non-free riders, given that free riders would be not influenced by the program to make a purchase decision and therefore would purchase the "exact same measure" without the rebate. But examining these results tell a different story. In 2017, participants were split between those who initially said they were influenced by the rebate (yes) and those who were not initially influenced by the rebate (no). The detailed analysis showed that eight initial "Yes" respondents provided a *contradictory response* to this question by providing a rating of "10." In contrast, eight of the original "No" respondents remained firm in their findings by answering "10" along with 11 additional program participants.

These contradictory findings continued in the 2020 evaluation. Examining these results by response category shows that the responses were split between those who initially said they were influenced by the rebate (yes) and those who were not initially influenced by the rebate (no). Four initial "Yes" respondents provided a *contradictory r*esponse to this question by providing a rating of "10." Similarly, three of the original "No" respondents remained firm in their findings by answering "10."

As Table 4 shows, ten respondents provided consistent answers in 2017, and four did in 2020. Overall, these inconsistent and contradictory results could have influenced the overall free ridership scores by as much as 36 percent in 2017 and 12 percent in 2020.

Table 4: Illustration of Contradictory Findings Regarding Purchasing Exact Same Measure

Source		Question: *Without the program, what is the likelihood that you would have purchases the exact same measure?"*	
Spire Rebate Application	*Did Rebate Influence Purchase Decision*?	2017 (n=28)	2020 (n=33)
	Yes	8	4
	No	***8***	***3***
	Blank	11	6
Spire Rebate Application	Yes	***2***	***1***
	No	0	0
	Blank	3	1
Contradictory Findings		10	4
% of Contradictory Responses		***36%***	***12%***

Additional Analysis

The evaluation team used this approach to examine the other questions relating to free ridership that were both on the rebate application and the customer survey. For example, the evaluation team received responses regarding the question, "*Had you already decided to purchase the equipment,"* from the customer to survey, indicating if the rebate influenced their purchase decision. The assumption is that a program rebate would not influence free-riders. However, the 2017 results revealed that 21 customers who initially said the rebate *did influence* their decision then c*ontradicted themselves and said they had also already decided to purchase the equipment.* Another nine customers who initially said the rebate had no influence confirmed that they had already decided to purchase the equipment.

There were also contradictory results in 2020. Specifically, 11 customers who initially said the rebate ***did influence*** their decision then ***contradicted*** themselves and said they had also already decided to purchase the equipment.

Conclusions and Recommendations

Table 5 summarizes these qualitative free ridership estimates based on this multi-year analysis for 2017 and 2020. Overall free ridership for 2017 was estimated at 36 percent. Most telling, *none* of these results suggest a free ridership higher than 44 percent, which differs significantly from the NTG algorithm used to calculate free ridership only from the customer surveys.

The free ridership rate dropped to 28 percent in the 2020 program evaluation, using this multi-question approach. The customer survey determined that free ridership was 45 percent. However, the analysis of the rebate application questions yielded a much lower rate due to the respondents "flipping" their answers. To arrive at a reasonable free ridership estimate, the evaluation team calculated the weighted average of the responses from two sources of data: the rebate program databases, not accounting for blank applications, and the weighted averages of the survey responses. This free ridership rate estimate incorporates data from both the rebate applications and the customer surveys. It is also slightly lower than the previously calculated free ridership estimate of 36 percent but is consistent with program planning estimates.

Table 5: Summary of Estimated Free Ridership Rates from Comparison Analysis

Question	2017 Free Ridership Estimate	2020 Free Ridership Estimate
No Rebate Influence by purchased exact same measure- Rating "8-10" – no and blanks	44%	14%
No Rebate Influence by Influence in Database- "1-3" Rating for No and Blanks-	29%	15%
No Rebate Influence by Yes Decided to Purchase	44%	15%
Revised Free Ridership Estimates from All Sources	36%	28%

The key takeaways from this multi-year analysis are:

- Determining free ridership is a complicated task and requires multiple approaches to ensure a consistent and reliable estimate.

- Program participant responses are not always reliable. One of the most surprising findings from these analyses was identifying how often customers changed their answers over time. The analysis revealed that a small percentage of these customers actually contradicted themselves by providing completely different answers to these questions at each point in time.

- It is important to monitor free ridership rates throughout the program cycle by including identical questions on the initial rebate application and the follow-up customer surveys. This additional layer of analysis will provide more significant insights regarding the overall reliability of free ridership estimates and provide another important data source in analyzing free ridership rates over time.

References

Illinois Technical Reference Manual, Version 6, Volume 4, Figure 4-6, p. 70

Johnson, Katherine 2017. "Appendix G: Memo on Free Ridership for Laclede Gas Company's Program Evaluation Results," *Energy Efficiency Program Evaluation Results for SPIRE MOE (Laclede Gas) Residential Water Heating and Space Heating Program and C&I Rebate and Custom Programs*, Prepared for Mr. Shaylyn Dean, Manager, Energy Efficiency Program. Prepared for Mr. Shaylyn Dean, Manager, Energy Efficiency Program. November 30.

_____. "Appendix D: Memo on Free Ridership for Laclede Gas Company's Program Evaluation Results," *Energy Efficiency Program Evaluation Results for SPIRE MOE (Laclede Gas) Residential Water Heating and Space Heating Program and C&I Rebate and Custom Programs*, Prepared for Mr. Shaylyn Dean, Manager, Energy Efficiency Program. November 6.

_____ 2017 and 2020, "Residential Participating Customer Survey- Residential High efficiency space Heating and Water Heating Program," Final, 9-22.