Evaluation of Peak Time Rebate (PTR) Baseline Methodologies

PREPARED BY:

CLEARSPRING ENERGY ADVISORS

CONTACT: STEVE FENRICK

STEVE.FENRICK@CLEARSPRINGENERGY.COM

Clearspring Energy Advisors LLC

Table of Contents

Introduction

Peak time rebate (PTR) programs are growing in popularity around the country. It is easy to see why. The program offers payments (or bill credits) to electricity consumers as a reward for helping their utility reduce wholesale power costs. The program is voluntary, and consumers maintain control of when (or if) they modify their electricity use in response to a peak alert signal sent to them via text, e-mail, or social media. Retail rates can remain unchanged and consumers on the program cannot be financially harmed, only rewarded, if they help the utility reduce costs. PTR programs typically see participation rates over 20% with high program satisfaction and low drop-off rates.

> ***"Capital Electric's PTR program has been a great program to engage and partner with our members to reduce their electric bills."***
>
> *-Paul Fitterer, General Manager of Capital Electric Cooperative*

There are a couple of critical items required to run a successful PTR program. First, precision in calling peak alerts is vital to the financial viability of the program. Rebate costs will add up rapidly and participants will get fatigued if too many events are called. Or wholesale or capacity cost savings will be lost if events are not called when wholesale billing peaks or system peaks occur.[1]

The second crucial aspect of a successful PTR program is calculating how much participants reduce their usage during peak alert events. Accurately determining the consumer baseline (i.e., what the consumer was expected to use absent a peak alert event) is imperative to program success. If people reduce their usage during events (e.g., run the electric clothes dryer after the event) they will expect to be paid for their efforts. Conversely, if people are paid rebates despite not changing their behavior, this will dampen future incentives to actively participate in saving the utility money.

In this paper we are going to examine two popular baseline methods:

1. A day matching method using the top three usage days out of the last ten similar days. "Similar" means the same type of day is used. For example, if the event occurs on a weekday, then the last ten non-holiday/non-event weekdays are used. We will call this the **3 of 10 Method**.

2. A regression method that correlates weather variables and day of week variables to each household's usage. We will call this the **Clearspring Econometric Method**.

[1] Clearspring Energy staff have extensive experience in calling peak alert events for PTRs and other load management programs. The paper will not be going into depth on that topic. However, please contact us if you have an interest in seeing how we might help with load management event dispatching.

To evaluate the two methods, we used real AMI hourly interval data from Capital Electric Cooperative (CEC) located in Bismarck, North Dakota. CEC has a fully deployed PTR program comprised of over 5,000 members.[2]

CEC provided interval data for 1,000 households from May 2018 through September 2018. We excluded from the analysis any PTR events or other load management events so that the interval data was for normal, non-event days. We then chose to simulate three peak alert events in each month (15 total for the 5-month summer season) and assumed peak alert events from 3 pm to 6 pm. This resulted in 15 simulated days and 45 simulated event hours. Since the actual usage was not actually a peak alert event, the actual data is the "true" baseline. This is what the participants would have consumed if no peak alert event were called. This is the very thing the PTR baselines are attempting to estimate. This provides the ability to accurately test each baseline method to see how it compares to the "true" baseline.

Statistical Results of the Two Baseline Methods

We examined two statistical results. The first is bias. Bias is determined by calculating if the baseline estimate is, on average, above or below the actual usage during events. If the baseline has a positive bias, then people will tend to be over-paid. If the baseline has a negative bias, then people will tend to be underpaid. Ideally, there is no bias so people, on average, are correctly paid. The second test is one of accuracy. Accuracy looks at everyone's baseline and calculates how the estimated baseline varies from the "true" baseline. The higher the average variance means that more people will be incorrectly paid (or underpaid) and by higher amounts.

The relationship between bias and accuracy is important. One could have a model with zero bias but one that still pays everyone vastly wrong amounts. The preferred baseline method has very little bias but high accuracy.

How do the two methods stack up regarding bias and accuracy? For the 1,000 households examined on CEC's system during the 45 simulated peak alert hours, the average usage of a household is 141.1 kWh. The 3 of 10 Method, on average, predicts usage at 160.6 kWh. For the same peak alert hours, the Clearspring Econometric Method predicts usage at 138.6 kWh. The 3 of 10 Method has an upward bias of 19.5 kWh (+13.8%) and the Clearspring Econometric Method has a downward bias of -2.5 kWh (-1.7%).

The bias of the 3 of 10 Method will also have implications on the estimated program benefits. If the utility is using the 3 of 10 Method they will erroneously inflate their estimated impacts by 19.5 kWh per participant for the entire summer. Since this is over 45 event hours, this equates to a kW impact of 0.43 even if no one actually responds to the peak alert. In other words, in each month the utility will think

[2] Clearspring Energy will be conducting the PTR event dispatching and rebate calculations using the econometric method for CEC.

the PTR program's impact is 0.43 kW higher than it truly is. If the demand charge is $20/kW, the utility will believe it is saving $8.60 more per participant ($8,600 per thousand participants) than it truly is saving. This will not only result in paying far higher rebates to participants for "phantom" wholesale savings but could also lead to non-optimal load management decisions that ultimately cost consumers substantial money and lost opportunities.

The second statistical measure to evaluate the two methods is the accuracy of the individual baselines. We measure this by taking the absolute average error for each of the 1,000 households examined. The error is the difference between what the method estimates the baseline, and the actual measured usage. We take the absolute error because we want to measure accuracy. We do not want a large positive error to cancel out a large negative error.

For the 45 event hours, on average, the 3 of 10 Method is either 21.6 kWh (15.3%) higher or lower than each individual's actual usage. This compares to the Clearspring Econometric Method which has an absolute error average of 5.8 kWh (4.1%).

Analysis For Entire Summer Season					
	Average Total Individual Usage in 45 Event Hours (kWh)	Average Bias over 45 Event Hours (kWh)	% Bias	Average Absolute Error over 45 Event Hours	% Absolute Error
Clearspring Econometric Method	**141.1**	**-2.5**	**-1.7%**	**5.8**	**4.1%**
Day Matching (3 of 10)	**141.1**	**19.5**	**13.8%**	**21.6**	**15.3%**

Based on the statistical measures of bias and accuracy, the Clearspring Econometric Method far outperforms the 3 of 10 Method. It also has other important attributes that are difficult to quantify. The econometric method cannot be "gamed", whereas the 3 of 10 Method could easily be manipulated by a participant that notices upcoming extreme weather and ramps up their usage, and thus baseline, for three days prior to an event. This is also why any sort of day-of "usage adjustment" should not be used in the baseline method. Additionally, we are of the belief that if the baseline is not accurately calculated, PTR program impacts will diminish over time as participant payments do not fully reflect the actions they are taking for the utility. The 3 of 10 Method will have far more inaccurate payments and this will tend to erode participant confidence and impacts as time passes.

Why does the econometric method perform better?

Two considerations arise when analyzing the 3 of 10 and econometric methods. One is what we term "**behavior variance**" and the other is "**weather sensitivity**".

Behavior variance is the natural ups and downs of electricity use at the household level. One day a consumer might have guests over and their usage is higher than normal, whereas the next day they might be out of town and usage is lower. Assuming the person does not react to PTR events at all, on that first day their usage will likely be above the baseline (zero rebate) but on the second day their usage will likely be below the baseline (positive rebate). If the rebates are paid after every event, the person will be paid a rebate amount despite making no intentional effort to reduce usage. However, if the two days are averaged (their usage and baselines) then they might not receive a rebate or will receive a lower rebate.

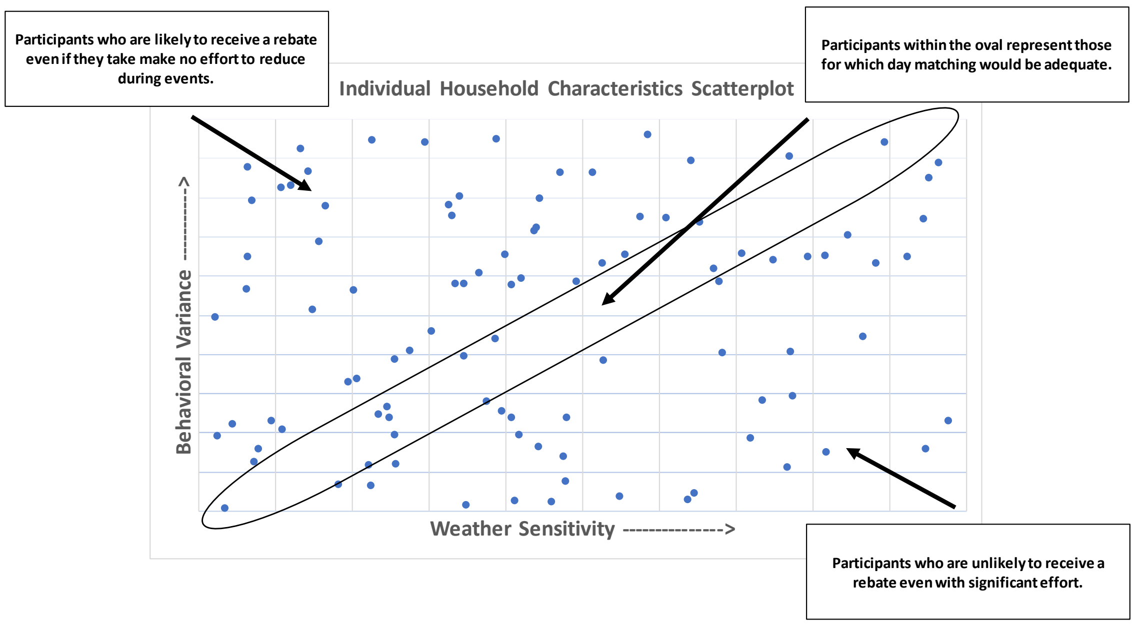
Weather sensitivity quantifies how a participant reacts to different weather situations. Generally, a participant will use more electricity in more severe weather, but the exact patterns and amounts will vary greatly from individual to individual. Some households may turn on their air conditioner when outdoor temperatures hit 75 degrees while others may not turn it on until the temperature reaches 90 degrees. Some houses may have single room air conditioners while others have central AC or no AC at all. Air conditioner sizes and building envelope efficiencies will be different from household to household resulting in drawing different amounts of electricity. In winter months, weather sensitivities from one household to another can vary even greater with differences in heating fuel types. While we did not test winter months in this study, we suspect the inaccuracies of the 3 of 10 Method at the individual level are similar in the winter due to the variety of household fuel types.

Everyone would agree that peak events typically happen during the more extreme weather patterns since there is a clearly observed correlation between weather and system wide electricity usage. Because of this, to accurately establish an individual's baseline, it is not adequate to simply take an average of prior usage for most households. The econometric method quantifies what specific weather sensitivity adjustment is needed for each individual to provide an estimation of their baseline usage customized to each event. The 3 of 10 Method makes no attempt to adjust for weather sensitivity either at the specific household or for the specific event day. Instead it relies on behavior variance to offset the effects of weather sensitivity.

While the effects of weather sensitivity and behavior variance are typically of opposite sign, and work towards cancelling each other out, there is nothing to suggest they are equal. Additionally, they certainly are not equal for every household and for every event. In fact, no single method that relies on a behavior variance adjustment to offset weather sensitivity will work well for every event or for every household because each event and household requires a different weather sensitivity adjustment. Some events will have very similar weather to the 10 days proceeding it and will only require a small adjustment, while other events will have very different weather conditions to the 10 days proceeding it and will require a much larger adjustment. The 3 of 10 Method assumes all households and events are the same, whereas, the econometric method customizes the calculation specific to the weather sensitivity of each household and adjusts for the type of event.

To see the impacts the 3 of 10 Method has on an individual level, reference the scatter plot shown below. The 3 of 10 Method, or any other method that relies on behavioral variance to offset weather sensitivity, only performs well for those households where both are equal (the households that fall within the black oval). Houses on the bottom right on the scatter plot have a relatively high amount of weather sensitivity and little behavioral variance. Perhaps they have a large AC unit but the temperature settings are preset and their schedule varies little day to day. Even if these households make a considerable effort to reduce usage they are likely to receive little or no rebate and will probably get discouraged over time. Alternatively, households located on the upper left of the graph have low weather sensitivity and a large amount of behavioral variance. Perhaps these households don't have an air conditioner, or it is a winter season and they do not have electric heat, but their daily schedules and activities vary from one day to the next. Households such as this will consistently earn rebates regardless of taking any action to reduce usage during an event, increasing the utility rebate costs with no corresponding wholesale benefits.

Illustrative Example of Where Inaccuracies Arise with 3 of 10 Method



Financial Implications of the Two Baseline Methods

In examining the financial implications of the two baseline methods, we simulated what participants would typically reduce electricity by during peak alert events. For 20% of the participants we subtracted 30% of their usage (this assumes around a 1 kW impact per event hour for 20% of the participants), another 20% of the participants received a 20% reduction, and then 10% received a 10% reduction. Notice that half of the participants are assumed to not react at all during PTR events. This aligns with our findings in conducting the calculations throughout the years. In our simulation, on average, we are assuming a 0.31 kW impact per participant. Our experience shows that an average impact per PTR participant is between 0.2 kW to 0.4 kW, depending on program specifications.

Since the underlying data is the "true" baseline and we know exactly what we subtracted to simulate the PTR program, we know both the "true" baseline for each participant and the exact impact. Thus, we know exactly what each participant should be paid in rebates and what the kW savings are to the utility in this simulation.

For simplicity sake, let's assume the utility has a monthly demand charge of $20/kW based on a one-hour coincident peak.[3] We will also assume the utility is paying $1 per kWh reduced as the rebate amount and that five monthly peaks were hit in the 45 peak alert hours.

The benefits are easy to quantify and will be the same for both methods (assuming inaccuracy has not eroded impacts). Across the 1,000 participants, the utility's **summer wholesale power savings will be $31,000** ($20/kW-month*0.31 kW per participant*1,000 participants*5 months).

What will the rebate costs, overpayments, and underpayments be? We can know this exactly because of our simulation and both methods are using the exact same data with the exact same impacts.

The 3 of 10 Method will drastically increase rebate costs for the utility due to the bias in the method and the inaccuracy at the individual level. Both methods will have error attached with them and will improperly pay rebates of varying degrees. Some of these will be "overpayments", paying participants money for reductions that did not actually occur. Some will be "underpayments", which is not paying participants or not paying them enough despite them making reductions.

The table below shows what the actual rebates that were earned by the participants (which are the same for each method) and then the overpayments, underpayments, and total rebates paid by the utility.

Rebate Costs With Identical Impacts For Each Baseline Method

	Participant Earned Rebate Amount	Overpayments	Underpayments	Total Rebate Cost
Clearspring Econometric Method	**$14,824**	**$1,694**	**$2,152**	**$14,366**
Day Matching (3 of 10)	**$14,824**	**$20,522**	**$506**	**$34,840**

As the numbers above show, if the utility had used the 3 of 10 Method the rebate costs would have exceeded the wholesale savings for the summer and the utility would have lost money on the program. This is due to the overpayments to participants that either did not reduce or did not reduce by as much as they were paid. To make matters worse, we calculated how much of the overpayments ended up in the hands of participants that made no changes to usage. **Half of the overpayments (over $10,000 for the 3 of 10 Method) go to participants that made zero impact on the coincident peak**.

In this example, keep in mind that if a utility actually used the 3 of 10 Method they might not be aware of the fact that their rebate costs exceed the power supply savings. Because of the bias, the utility might believe the kW impact per participant was 0.75 kW rather than the 0.31 kW that we know occurred due to our simulation.

---

[3] The financial benefits of PTR can be substantial with varying wholesale power arrangements. However, they will vary based on the wholesale rate and the ability to precisely call peak alerts. Clearspring can provide your utility with a cost/benefit estimate if you are interested in finding out if a PTR program makes sense at your utility.

The 3 of 10 Method could be modified to reduce the bias. Perhaps the bias would be reduced if the method had been 3 of 5 instead of 3 of 10. However, due to the much larger inaccuracy of the day matching method, in general, this would have pushed the method's underpayments substantially higher. The day matching method should be chosen carefully, and the utility should be aware of the inaccuracy problem regardless of the day matching method that is chosen.

A Word on When to Do the Baseline Calculations

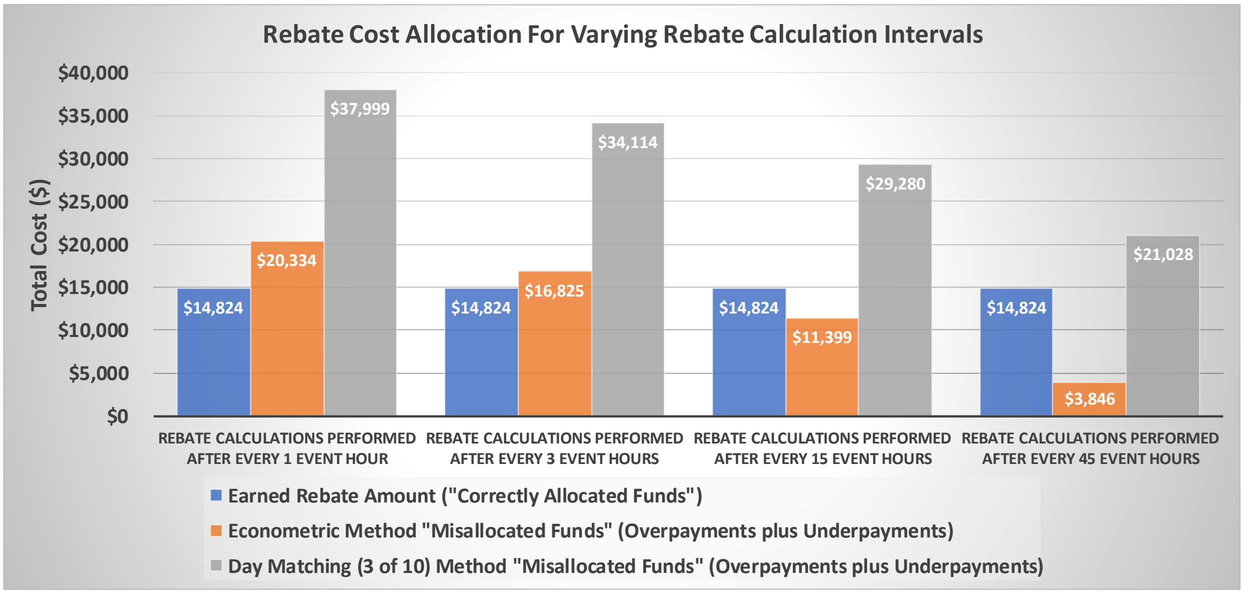
The issue of when to do the PTR baseline calculations and give rebates to participants is an important program consideration. Giving near real-time feedback to participants regarding their impacts and rebate amounts provides participants with greater awareness, increased customer engagement, and more immediate gratification for their efforts. However, there is an important consideration to factor in regardless of the baseline method being used. The fewer the number of events that are aggregated between calculations, more error will exist, and more money will be misallocated.

How much does this matter? The following table displays the overpayments and underpayments for the entire summer season of each method for different aggregation (or calculation) periods. [4] Again, due to "behavior variance" a utility measuring impacts after every event may think it has higher impacts than what it is truly saving.

	Overpayments			**Underpayments**	
Event Hours	Clearspring Econometric Method	Day Matching (3 of 10)		Clearspring Econometric Method	Day Matching (3 of 10)
1	$14,347	$33,494		$5,987	$4,505
3	$11,358	$30,326		$5,467	$3,788
15	$7,079	$26,548		$4,320	$2,732
45	$1,694	$20,522		$2,152	$506

[4] If calculations are made after every event (3 hours) there will be 15 calculations for the summer. If the calculations are done after 45 hours, then there will be 1 calculation. The more calculations and the fewer hours that they comprise, the larger the potential is for participants to "bank" rebates even with no reactions.

The following graph compares the earned rebate amounts for the summer ($14,824) to the amount of misallocated funds for each method and for each calculation period. Misallocated funds are the sum of the absolute value of overpayments and the absolute value of underpayments.



Conclusion

The Clearspring Econometric Method has far less bias and is more accurate than the 3 of 10 Method when comparing to actual data from 1,000 households on Capital Electric Cooperative's system during the summer of 2018. The day matching method has the advantage of being more understandable and transparent to participants, however, it does have the downside of being vulnerable to participant gaming.

Using our simulation of PTR events, the financial performance of the program was far superior using the econometric method. Overpayments of rebates were drastically lower compared to the day matching method and underpayments were small with both methods. The rebate costs for the same exact impacts using the day matching method were $19,000 higher for the summer months per 1,000 participants relative to the econometric method. Additionally, the 3 of 10 Method will artificially inflate per participant impact estimates and will mislead decision-makers into thinking system demands have been reduced far more than what truly occurred.

The rebate costs and inaccuracy of calculating the rebates after every event are considerably higher compared to waiting for the summer to end and the participants will receive larger rebate checks, which may motivate continued participation. The advantage of waiting needs to be weighted against the advantage of giving participants quicker feedback.

Our conclusions are consistent with a similar but more comprehensive study conducted by Freeman, Sullivan, & Co. on San Diego Gas and Electric's PTR program. For additional reading on this topic, that report can be found at http://www.calmac.org/publications/SDGE_PTR_Baseline_Evaluation_Report_-_Final.pdf.

If your utility is considering how to lower wholesale bills with low upfront costs, lower retail electric bills, and engage consumers a PTR program may be a great opportunity. Our team at Clearspring Energy has years of experience helping clients with PTR programs, rate design (including demand rates and interruptible rates), and other load management programs. We would love the opportunity to discuss if a PTR program makes business sense and how to best design it for your utility's circumstances.

Clearspring PTR Services

Clearspring Energy staff are experienced in helping utilities consider, test, and deploy PTR and other load management programs. We can help your utility determine if a PTR program makes economic sense and will work based on your power supply, AMI data, and your existing load management portfolio. If the program does have economic benefits, our staff can help move it from the drawing board to full deployment. We can offer the following PTR services:

- **Feasibility** of adding a PTR program from a financial, data, and portfolio perspective. This typically includes a quick examination of your AMI interval data to assure it can support a PTR program. In this we can provide recommended program specifications.

- Setting up statistically valid **pilots** to validate impacts, set program rules, receive participant feedback, and gain staff knowledge prior to full deployment.

- We can examine which **day matching method** would be best for your system and how that compares to the econometric baseline method.

- We can **run the rebate calculations** for your PTR program using either the econometric or day matching method on whatever period you choose (after every event, monthly, seasonal, or annual).

- Our staff are experts at **calling and dispatching peak alert events**. We have developed software algorithms and tools to precisely call PTR (and other load management) events to maximize program margins.

- Our team has helped utilities **recruit** participants (typically a 25% initial participation rate), educate participants and notify them of events through text and e-mail messaging.