Alexandra Czastkiewicz, EcoMetric Consulting

Presentation Title: How to evaluate Pay for Performance Programs: A Payday for Participants and Utilities

Abstract: The evaluation reviewed a new pay-for-performance (P4P), whole building DSM program offered to commercial and industrial customers in a major North American city. The program has been well received by participants due to the low amount of administrative burden and flexibility of installing capital projects or operational and maintenance changes. The proposed paper will delve into the benefits and drawbacks of increasingly popular P4P programs and the improvements made throughout our team's evaluation of the program's energy impacts and processes. The rebating process for the program is simple, a payout of $0.04 per kWh saved. This approach encourages participants to continually improve as they can collect rebates for up to four years after beginning program participation. In the program's first full year, the gross average program savings were 7.7%, with most sampled participants meeting the minimum 5% savings threshold. Another program requirement is a full year of daily utility data for the baseline and post retrofit period. When comparing the savings for the same project a monthly billing regression resulted in annual savings of 13.1%, while the daily regression led to a savings of 14.6%. These program requirements can leave fewer savings on the table and provide a high level of statistical accuracy. Traditional methods of calculating measure level savings often include many assumptions and cannot easily account for operational differences, weather changes, or production changes. Billing regressions normalize utility bills with typical meteorological year data, and because the regression is done using actual site utility data it captures and accounts for operational changes that might not be captured or accounted for in traditional methods of savings calculations which might over or underestimate savings. In the next evaluation year, our team is planning to improve the daily utility data model to utilize hourly utility data for billing regressions for a more robust analysis of energy savings. Hourly data would allow the evaluation team to leverage machine learning algorithms to better understand what is driving savings at the building-level. Further, our team discovered that a holiday model used to calculate savings on holidays for the program's main participant was statistically insignificant. We recommended that the savings be calculated by leveraging one model with a dummy variable to identify holidays and reduce their burden of creating the holiday models in their next par-for-performance period. The program could also be improved by setting a minimum p-value threshold for the regression models to ensure only significant models are being used to calculate savings. With open communication between the technical reviewer and evaluation team, it can be assured accurate savings are being calculated before they are reported. As our team continues to evaluate the popular and growing program, our goal is to retain the low administrative burden on participants that makes the program so popular while also leveraging project data to conduct a robust impact and process evaluation. The audience will learn what makes this P4P program so successful, how evaluations can provide valuable insights into savings drivers and how this success can be replicated in their jurisdiction.