THE E2E PROJECT'S EDUCATION PROGRAM

CAUSE AND EFFECT: USING RCTS AND MACHINE LEARNING TO IDENTIFY INTERVENTIONS THAT WORK

BECC WORKSHOP, SACRAMENTO, CALIFORNIA

OCTOBER 15, 2017

Part 1: Randomized Controlled Trials

Instructor: Professor Maximilian Auffhammer (UC Berkeley)

Randomized control trials (RCTs) are the gold standard for evaluating the effectiveness of programs in a wide variety of fields, including medicine, education, and international development. This course will give an overview of RCTs and describe how they can be applied to evaluate energy-efficiency programs. The course will also provide practical suggestions on how RCTs can be implemented and alternative approaches if an RCT is infeasible. The material covered is designed for energy-efficiency professionals with some training in statistics.

Introduction to Randomized Control Trial Research Designs

Topics: Energy-efficiency program evaluators face a difficult task. In order to identify savings from a program, they need to describe how much energy would have been consumed if the program hadn't existed. In other words, they need to provide a "counterfactual" description of the world. This session will describe common approaches to developing counterfactuals, including both observational studies and randomized control trials. We will discuss inherent challenges, including sample selection, statistical biases from omitted variables, and external validity.

Implementing Randomized Control Trials to Evaluate Energy-Efficiency Programs

Topics: A randomized control trial can be designed in many ways, but the treatment and control groups should be identified prior to implementation of the program. This session will describe common issues confronted in developing randomized control trials including identification of the treatment and control groups and the importance of random assignment to these groups. Good treatment and control groups enable an evaluator to pinpoint the impact of a program and eliminate other influences.

While randomized control trials offer an effective way to measure the impacts of a program or policy intervention, it is often neither practical nor appropriate to mandate or force a group of consumers to receive a "treatment." Fortunately, there are experimental research design alternatives that do not require mandatory assignment. This module will introduce some of these alternatives (including randomized encouragement designs, recruit-and-delay, recruit-and-deny).

Quasi-Experimental Approaches

Topics: In some empirical contexts, randomized control trials are simply not practical and/or infeasible. In those cases, quasi-experimental research designs can be an effective substitute. Quasi-experimental studies assign households to treatment and control groups by a method other than random assignment. Their effectiveness depends on program details and the data available. This session describes a set of quasi-experimental approaches, provides examples from the energy efficiency context and discusses their limitations. We will review examples of effective quasi-experimental studies.

Part 2: Machine Learning Methods

Instructor: Professor Mar Reguant (Northwestern University)

In cases where randomized controlled trials are not possible, we need to turn to alternative methods to quantify the savings from energy efficiency programs. Smart meter data creates new opportunities to evaluate energy efficiency programs, but also presents new challenges. Machine learning is an exciting new toolbox that can be used to leverage smart meter data. This course will provide an overview of machine learning and discuss how these methods can be used for energy efficiency evaluation. This material is designed for energy-efficiency professionals with some training in statistics.

Machine learning to evaluate energy efficiency projects

Topics: A central challenge in energy efficiency evaluation is figuring out what would have happened in the absence of an energy efficiency upgrade. Machine learning methods are a series of tools for data-driven prediction which are designed to predict well out of sample. In this session, we will provide an overview of existing machine learning methods, including “supervised” tools like LASSO and random forests and “unsupervised” tools like neural networks. We will then discuss how these methods can be used in the context of a specific energy efficiency evaluation, where we use machine learning to generate predictions about counterfactual energy consumption. We will discuss the benefits of using machine learning tools for energy efficiency evaluation, as well as the limitations of this method in a non-randomized setting.

Machine learning to better target energy efficiency projects

Topics: In many cases, energy efficiency projects will work well for some program participants and not for others. Understanding who will benefit from an upgrade can help design cost-effective policy. Because machine learning is designed for prediction, we can use these methods to identify “high-value” participants. In this session, we will discuss how machine learning can be used to understand which interventions are working, and who benefits from these interventions. Using this information, can we use machine learning-based targeting to improve the overall returns from interventions? We will discuss these issues using examples of machine learning applications in real-world energy efficiency interventions.