

ENERGY

Opting for the Unknown

Evaluation of Opt-in Behavior Programs

Ken Agnew

December 8, 2014

Outline

- Historical Context in energy program evaluation
 - Early years
 - The 2000s.
 - Last Five Years
- Highlights of the White paper.
- Recommendations – one year on
 - Example
- Conclusions

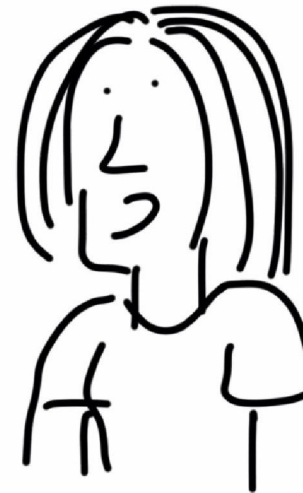
Very Funny . . .

What kind of evaluation
did you need?



Our 3 year project is coming
to an end and were told we
needed an evaluation.

What kind is that?



freshspectrum.com

The Early Years

- Quotation from a 1983 Article in Evaluation Review
 - *Because participation in these programs is voluntary, a self-selection bias may be present. Therefore, one must determine whether observed changes in energy use are due to the program itself rather than the composition of the group that accepts the program.*
- 1980s Mimi Goldberg and Ken Train write paper using double inverse Mill ratio to address self-selection bias.

My first 10 years . . .

- Self-selection, and the possibility of selection bias in results, strangely absent in reports and conference papers.
- Billing analysis (Consumption data analysis)
 - Tracking data for a program – List of participant who had installed an efficient furnace, Insulated etc.
- What would have occurred in the absence of the program
 - Pre-post difference
 - Weather-normalization
 - Comparison groups and/or pooled, fixed-effects models
- Free Ridership??
- Appropriate baselines??
- What does this have to do with Opt-in behavior programs?

The Last 4 years . . .

- Opportunity to do first evaluation of a program set-up with a Randomized Controlled Trial experimental design.
 - Eligible population randomly split (ideally within strata)
 - Best possible proxy counterfactual
 - Best we can do with respect to establishing causal link between treatment and effect.

- Relief after years of hand-waving.

- Provided a new lens on other impact evaluations

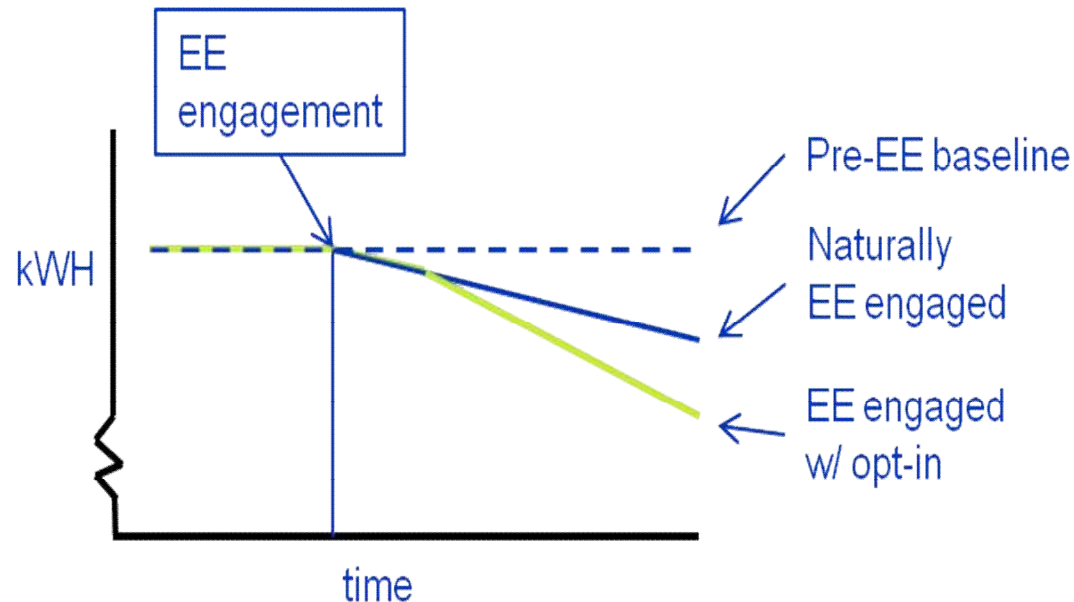
Opt-in Rebate programs from RCT Perspective

- Put comparison group in context
- Clearly illustrates the dangers of self-selection
- Clarifies free ridership

Highlights of the Whitepaper

- Primary goal was to illustrate the challenge of self-selection in estimating savings outside of an RCT experimental design.
- We illustrate the general challenges inherent in the consumption data analysis that we use to evaluate these kinds of program.
 - Kinds of non-program effects in pre-post analysis
 - Challenges of matching comparison groups
 - The limitations of pooled approaches to address concerns.
- Report recommends a combination of Variation in Adoption and match comparison approach.

Question of motivation



- Real question: How does opting-in to a web portal assist the already-engaged customer?
- The pre-engagement baseline assumes all change due to the portal.

Recommendations I would add . .

- Figure out, in advance, how to avoid VIA or matched comparison group approaches.
- Let's only use quasi-experimental approaches where absolutely necessary rather than as an easy alternative to good design.
- Going to require major changes in habits for (some) program folks who are used to dealing with evaluation after the fact.

Where is Energy Program Evaluation At Compared to Other Areas?

- Recent NYT article “Health Care Myths” (11-23-2014) About the Oregon Medicaid natural experiment.
- Studies using the randomized experiment yielded unexpected results that, contrary to anecdotal evidence, Medicaid was not associated with a drop in the use of emergency rooms by those that got received Medicaid.
- Of greater interest, RCT approaches are not as widely used in medicine as we might expect

In a recent review of major medical journals, (Finkelstein, MIT) found that more than 80 percent of published papers on medical interventions, like drugs and operations, involved randomized trials. Fewer than 20 percent of studies about how to actually practice medicine used randomized groups of patients or doctors.

Conclusions

- Self-selection has been present, and acknowledged, in energy program evaluation for decades.
- RCT experimental designs for behavior programs have provide a useful lens through which to better understanding the implications of this.
- The way forward with opt-in programs, and behavior programs in general, should not be the application of the well worn and limited methods that we have used for years.
- Instead, we need to design programs for evaluability by including randomization from the design phase.

- Innovations in the evaluation of behavior programs are actually innovations in the design of behavior programs – so that the evaluation is too simple to deserve a conference paper.

Opting for the Unknown

Evaluation of Opt-in Behavior Programs

Ken Agnew

ken.agnew@dnvgl.com

608-259-9152 x70207

www.dnvgl.com

SAFER, SMARTER, GREENER