

PROGRAM EVALUATION IN APPLIED ECONOMICS

BEHAVIOR, ENERGY, AND CLIMATE
CHANGE CONFERENCE

NOVEMBER 17TH, 2009

Hunt Allcott

MIT Department of Economics and [ideas42](#)

Overview: Program Evaluation in Applied Economics

□ **Two observations:**

1. There is a large recent shift in applied microeconomics toward using randomized field experiments for program evaluation
 1. Examples from energy efficiency and elsewhere to illustrate why randomized controlled trials (RCTs) are important
2. There has been substantial “learning-by-doing” in experimental design that makes RCTs more feasible
 1. Examples of innovative designs from health, international development, education, labor markets, and finance
 2. With an eye to applications in electric utilities

The Shift to Randomized Field Experiments

- New think tanks do randomized experiments in development, health, education, finance, etc.
 - ▣ MIT Poverty Action Lab, Yale Innovations for Poverty Action, Harvard ideas42, Analyst Institute, etc.
- Key difference from earlier academic experiments:
 - ▣ Applied, policy relevant questions
 - ▣ Partnerships with governments, NGOs, businesses
 - ▣ Development example: national to individual analyses
- Economists no longer believe analyses without experimental (“quasi-experimental”) randomization
 - ▣ Otherwise fail the peer review and publication process.

A Few Questions

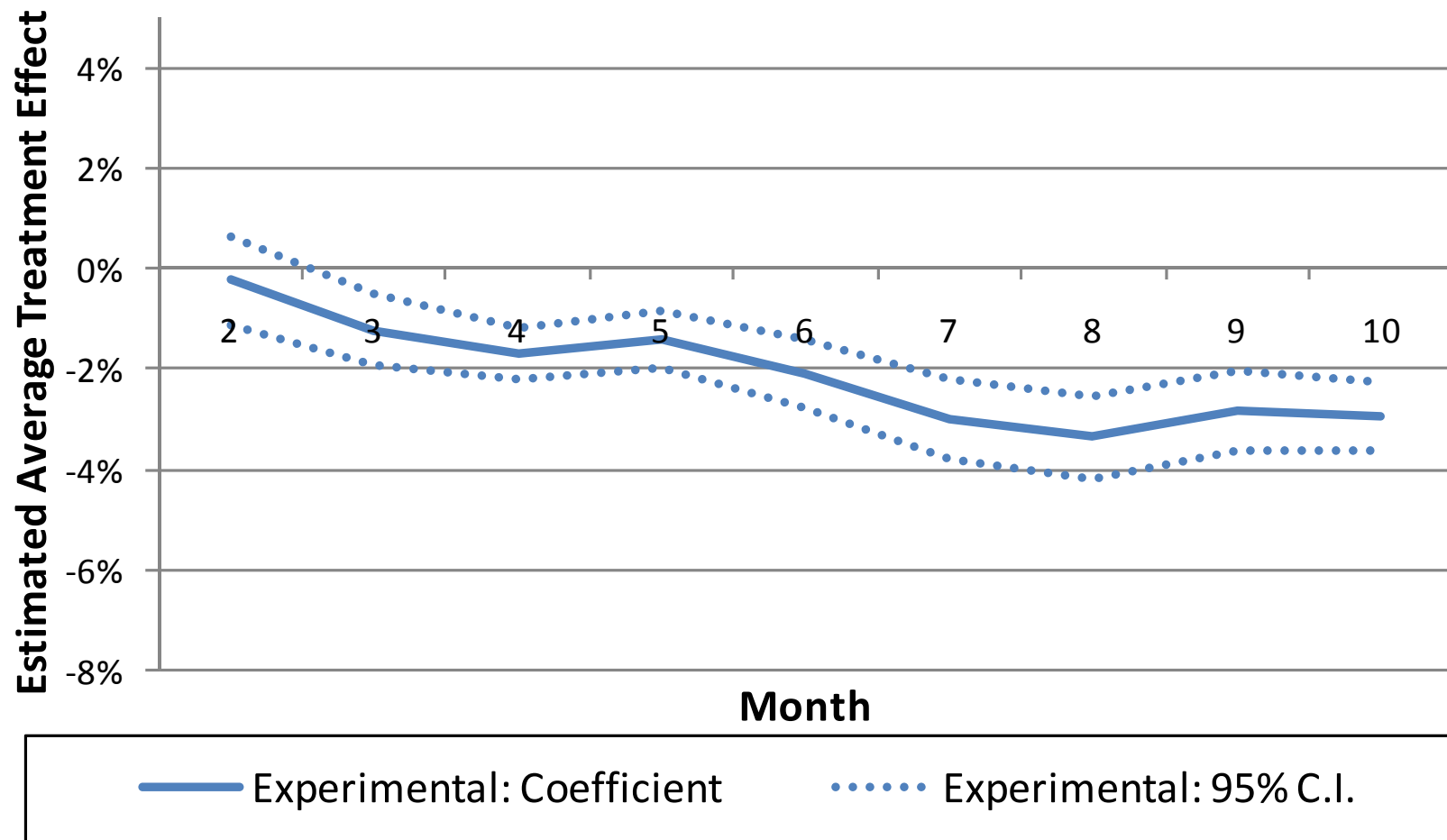
- Why should utilities/program managers care?
- How important are randomized controlled pilots for estimating a program's causal effects?
- Can you give a tangible example from an energy efficiency program of why this is important?

Example: OPOWER (Positive Energy)

- Home Energy Reports: Letters with neighbor comparisons and energy conservation tips
- Eligible households randomly assigned to receive report letters (Treatment) or not (Control).
- Question: What is the Average Treatment Effect of the OPOWER program, i.e. how much does it reduce electricity consumption in the average household?
- “Social Norms and Energy Conservation.” MIT Center for Energy and Environmental Policy Working Paper #2009-013. Under submission to the American Economic Journal: Applied Economics.

OPOWER Treatment Effects: Monthly Reports

2009 Treatment Effects by Month

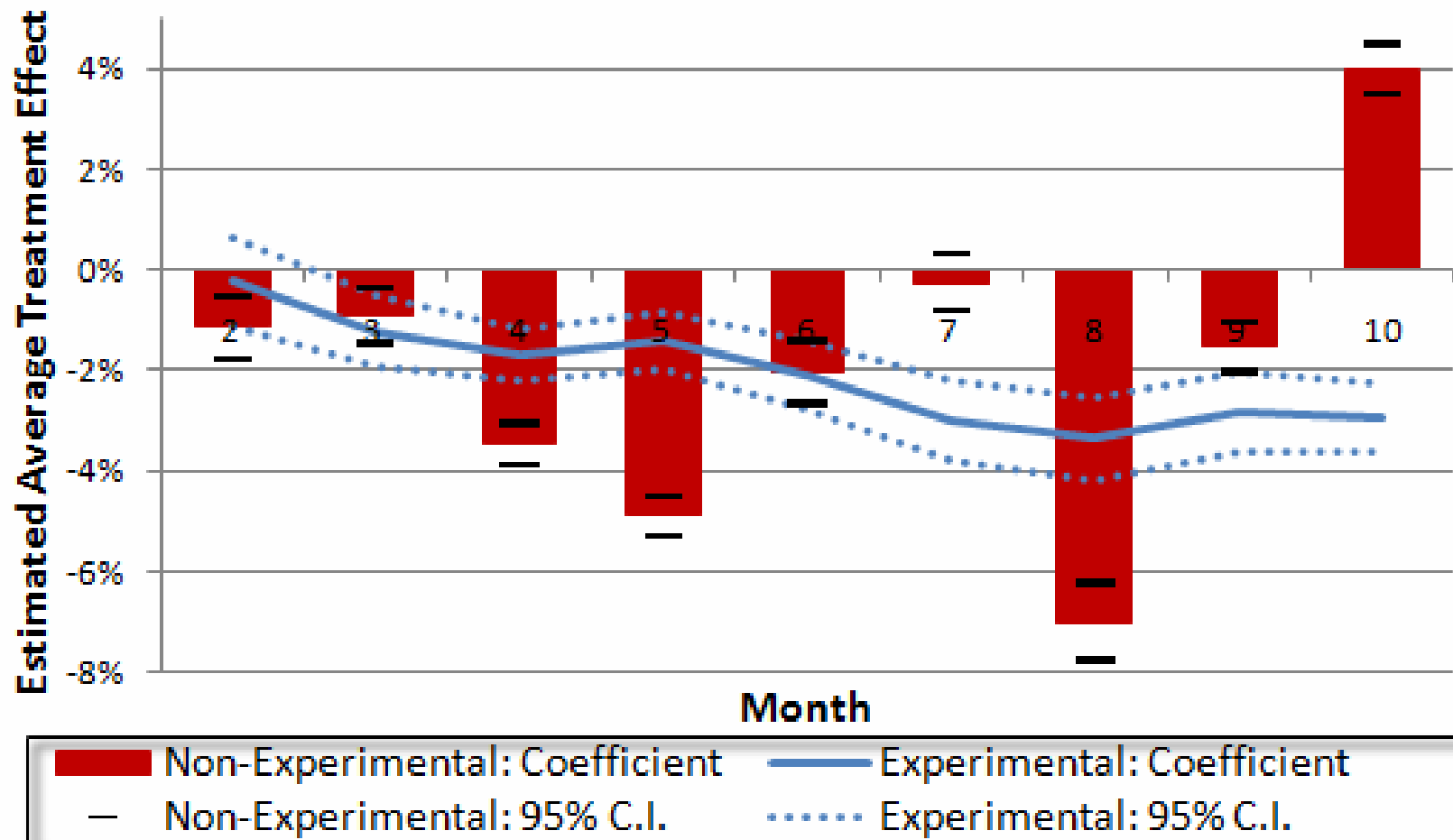


Is Non-Experimental Evaluation Possible?

- Could we have estimated the correct treatment effects without a randomized control group?
 - ▣ A common approach: look at participants' year-on-year savings and control for observable factors
 - Energy saving club at large North American utility
- Example with OPOWER data: let's estimate treatment effects with treatment group data only
 - ▣ Control for weather using 4th-degree polynomials in heating and cooling degree days
 - ▣ Control for month
 - ▣ Use fixed effects for all time-invariant characteristics

The Importance of a Randomized Control Group

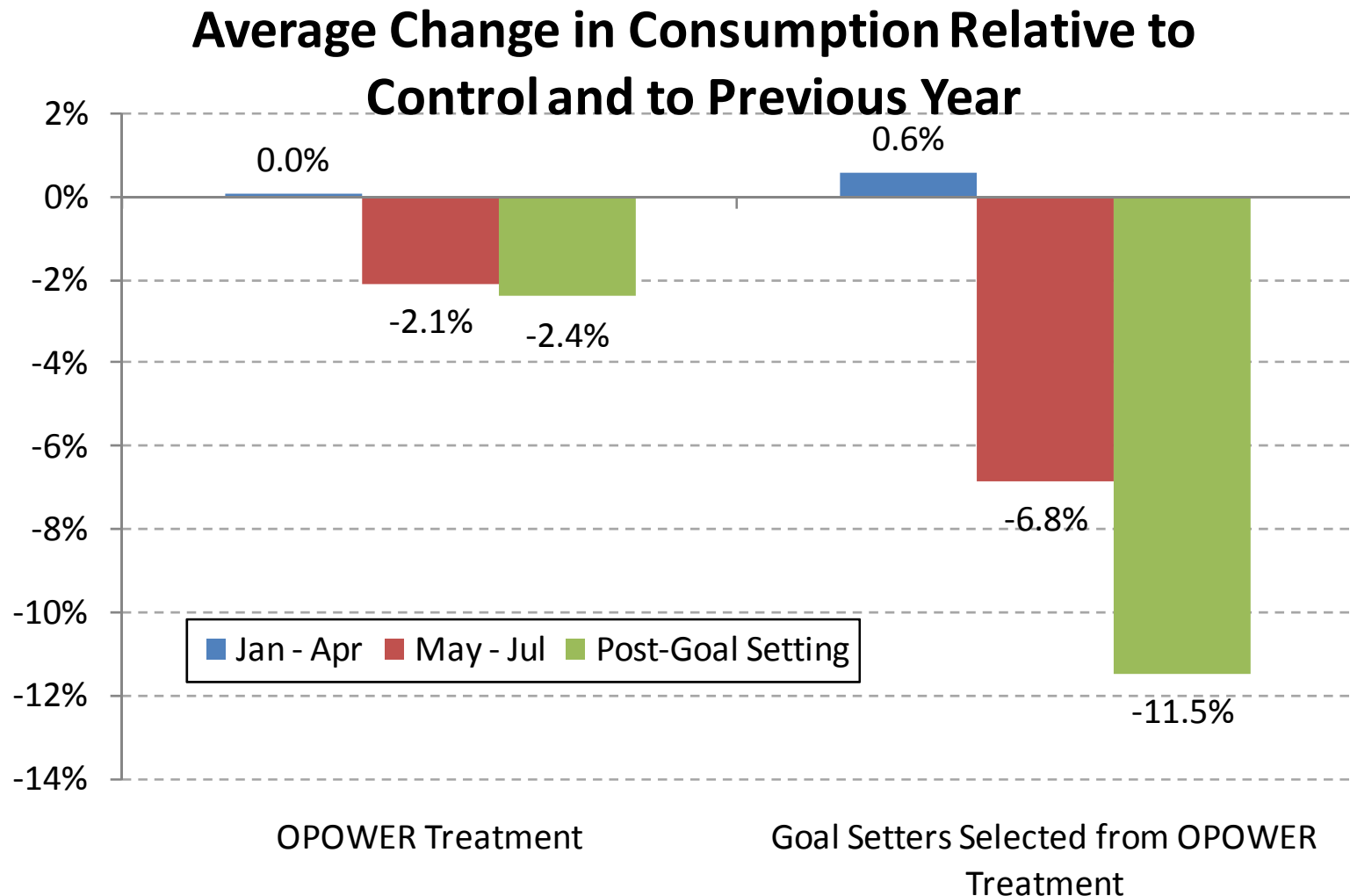
2009 Treatment Effects by Month



What Control Group to Use

- So we need a control group . . .
 - ▣ Weather and unobservable differences over time
- Who would be a good control group?
- In an opt-in program, can we compare enrollees to a control group from the general population?
 - ▣ Home energy use information displays at a large North American utility
 - ▣ Numerical example: OPOWER Goal-Setting program
 - OPOWER treatment group customers offered an optional goal-setting program

Selection into OPOWER Goal-Setting Experiment



The Selection Problem

- Example of a broader problem: selection bias.
 - ▣ Job training programs: compare wages for those who enrolled in the program to those who didn't?
 - ▣ College attendance: Compare the incomes of college grads to high school grads?
 - ▣ Economists no longer consider non-enrollees as good controls for enrollees, even after controlling for observable characteristics
 - Unobservable characteristics: motivation, ability
 - Energy efficiency: those that opt in are more “engaged”

Randomized Trials Increasingly Feasible

- So far: illustrated why randomized controlled trials important for meaningful (internally-valid) results
- For some this is old news
 - ▣ Electricity the first application of RCTs in social science!
- Randomized trials are used to evaluate some utility programs, but not nearly all. Why?
 - ▣ Is randomization easy?
 - ▣ What if we can't force into treatment or control?
 - ▣ What about *external* validity?
 - ▣ Is it fair?
- Examples of experimental design in other industries

“Points of Randomization”

- *Feasibility concern: Is randomization easy?*
- Letters or e-mails to businesses and households
 - ▣ Bertrand and Mullainathan (2004) resume experiment
 - ▣ Bertrand, et al (2010) bank loan experiment
 - ▣ Businesses: Capital One, OPOWER, etc
- Call centers
 - ▣ Analyst Institute (Rogers, et al) Get-Out-The-Vote experiments, implementation intentions
- Depends on the program being evaluated: Can also randomize by store, village, day, auditor, etc.

Randomized Encouragement

- *Feasibility concern: Situations where individuals can't be "forced" into or out of treatment or control.*
- Randomized encouragement
 - ▣ Job training program. Recruit people for a survey. Treatment group given participation incentive (cash).
 - ▣ Commitment savings program. Randomize *offer* of a different loan product
 - ▣ Treatment group has higher probability of enrollment
- "Encouragement" in energy efficiency programs
 - ▣ Different marketing approaches through letters, calls
 - ▣ Simultaneously test marketing and program effects

External Validity and Generalizability

- “Should we expand program X to a larger group?”
- Opt-in energy use monitors and RTP programs
 - Want: Conservation/elasticity in larger future program
 - Have: Conservation/elasticity for most engaged customers!
- Want experimental population similar to future
 - Wolak (2006) Anaheim CPP: concerted recruitment
 - Incentives such that take-up \approx future take-up
- Replication in other locations and populations
 - Innovations for Poverty Action, early TOU experiments

Are Randomized Trials Fair?

- *Feasibility concern: Is randomization fair?*
 - ▣ Are PUCs (and utilities) willing to allow different treatment to different customers?
- Fundamental question across domains
 - ▣ Clinical trials: some get medicine and some don't
 - ▣ Poverty alleviation programs: some students get new books and some don't

Are Randomized Trials Fair?

- Time constraints in implementation: phase-ins
 - ▣ Kenya de-worming medicine experiment
 - ▣ PROGRESA welfare program in Mexico
- Resource constraints: lottery is perhaps more fair
 - ▣ Utility pilots are by definition limited in scale

When Randomized Experiments Work

- At least several hundred participants:
 - ▣ Residential and small commercial programs
- When outcome is clearly defined:
 - ▣ Energy savings, price elasticity
- When the outcome is uncertain
 - ▣ If we know that a program works, why test it?
- Unobservable characteristics or actions
 - ▣ “Free riders,” “rebound effects,” etc
- *Many pilot projects don't fit these criteria*
 - ▣ But perhaps more than we see today?

Conclusion

□ **Two takeaways:**

1. Randomized controlled field experiments are often necessary for meaningful evaluation of pilot programs
 1. OPOWER examples of a broader consensus among empiricists across a variety of domains
2. Randomized experiments are increasingly feasible
 1. Experience from other domains in experimental design and execution with businesses, NGOs, and governments

Businesses that Test Products and Services with Randomized Trials

- Consumer finance:
 - ▣ H&R Block, ICICI Bank, Capital One, Toronto-Dominion
- Restaurants:
 - ▣ Subway, Hardee's, Carl's Junior
- Internet
 - ▣ Amazon, Ebay, Google, Yahoo
- Harrah's Casino
- Davenport, Thomas (2009). "How to Design Smart Business Experiments."
Harvard Business Review.

Other Organizations that Use RCTs

- International development:
 - ▣ World Bank, IFC, International Child Support
 - ▣ Poverty Action Lab, Yale Innovations for Poverty Action
- Get-Out-The-Vote organizations:
 - ▣ Analyst Institute, AFL-CIO
- Federal Government:
 - ▣ HUD Moving to Opportunity
 - ▣ FDA clinical trials
- Banerjee, Abhijit, and Esther Duflo (2008). “The Experimental Approach to Development Economics.” MIT Working Paper.