EEN Forum 2009 Notes, 6/8/2009, 1 p.m.
By: Courtney Cameron

Session 3: Some Examples and Lessons Learned from Experimental and Quasi-Experimental Evaluation Methods (Room 309)

Paul J. Ferraro
Department of Economics
Andrew Young School of Policy Studies
Georgia State University

Experimental and Quasi-experimental Evaluation Designs Perspectives from Academia

Way TO ESTIMATE COUNTERFACTUAL OUTCOMES IS TO COLLECT data so that an actual program affect will be visibly different

- Region wide drought between states and led to a dispute between states and sectors
- Randomized experiment that targeted residential areas

MOTIVATION

- Professor went to multiple utilities to "sell" design
- Utilities interested but didn't want to do. Or said effectiveness was not important(didn't care about what they were doing, only did it because they had to)
- 1 utility hired a water efficiency coordinator. Boss only cared about money and how much it would cost, if anything. And if it counted toward their effort
- Water consumption declined in control group
- Water consumption declined in control group
- Strong social norms most effective (5% decline in consumption, on average)
- Can evaluate speed of impact and heterogeneity across household
- Social norm message is addressed to household, reminds them of the drought, provides how much water used, their consumption amount, "every drop counts". Prepared their consumption amount to their neighbors
- Can easily calculate cost effectiveness ($/avoided gallons consumed)

WHAT WAS DONE WITH INFO?

- Results were easily understood and credible
- Water consumption reductions without price increase causes revenue declines
- 20009 drought ended
- 200 calls were received from letter recipients and ½ were annoyed or angry about social comparison (<1%) { angry they were compared to other people in the county}
- Waterwise council wants to see follow up

REGUALTIONS AND POVERTY

- Hard to randomize regulation
- How effective are they? What are their socioeconomic impact?
- PROTECTED AREAD AND SOCIAL WELFAR
- Studies that look at condition of local people only tells us they are poor

QUASI EXPERIMENTAL APPROACH IN TILAND

- Geo referenced data to control confounders

- Secondary data (census, household surveys to measure poverty)
- Matching methods to identify counterfactual (assume protection uncorrelated with poverty conditional on matching covariates)
- Assess sensitivity of results to unobservable confounders
- MOTIVATION
- Interest to policy maker practitioners
- Difficult to answer in context single project or program
- Difficult to answer in short time
- Politically sensitive (bad for poverty, budget decrease)
- Demand for evaluation unlikely to come from implementers
- CONCLUSION
- Simple comparison of expost outcomes between areas with PAS and areas without PAS imply PAS exacerbate poverty
- Simple before-after comparison for areas with PAS suggest PAS strongly alleviate poverty
- Controlling for confounders No evidence that PAS exacerbate poverty
- USE?
- Has to published
- Have to look at heterogeneous treatment effects and dose response relationship (too much? Too little?)
- Look at mechanisms
- Look at other nations (country by country to get a picture of evidence base)
- ENVIRONMENTAL PERFORMANCE PAYMENTS
- Conditional cash transfers
- Thereat tot effectiveness include poor admin. Targeting and self selection by those planning on meeting performance targets
- COSTA RICA
- Compliance good
- Little administrative targeting (70% country priority) large #
- Self selection: 71% contract on land classified as unfit or with strong limitations for agriculture
- Secondary agriculture less likely to grow crops in 96', less likely to live on farm. More likely to have off farm income, more education, larger farms
- OTHER APPROACHES
- Uganda, randomized experimental design
- ECUADOR: attempt to create exogenous variation in treatment assignment in program scaling
- Poverty eligibility, random phase in,

QUESTIONS

1. Change in people who didn't et social norm message? Just tip sheet? No, real reduction in water consumption.
2. Difference in price between tip sheet and social norm message? Same, tip sheet was an extra sheet of paper. no real difference
3. Comparison same for all people. Comparing people to other people was less desirable

4. Control group had a reduction unrelated to messages. Was not included in overall outcome. Only extra
5. 8% for control group. 13% reduction overall.
6. What/how to get client to want info? CR doesn't necessarily have interest in info. How does evaluator get them interested? Must have incentives for people. And us government, people funding programs must make it pertinent for coordinators of programs so they have an incentive (their boss telling them) incentives must change so evaluators get programs to change.
7. How does a program being implemented in one state differ than a state that doesn't implement one? Daylight savings time. Indiana had counties that did/didn't take place. Showed difference in energy consumption

Lou Nadeau
Economist
Eastern Research Group

- talks about context of evaluation
- *Deborah rouggs? (relates to her speech)
- PURPOSE
- Talks about complex methods in program evaluation
- FOCUS on use of these methods in evaluations when the methods aren't well understood. Especially incases where the results show the program in ineffective or not meeting its objectives
- PEOPLE
- Program managers, evaluators provide an objective answer to evaluator questions, methods
- WILL METHOD WORK?
- Evaluator, can the method be applied to the available data generate a valid estimate of the program impact?
- Program manager, will the method show my program is successful?
- COMPLEX METHOD? WHAT ARE THEY?
- Often involve advanced statistical technique limited understanding by program managers and evaluators become the sole expert
- Specific techniques regression analysis adjusting for selectivity. Propensity score matching
- WHEN DO U USE THEM
- Prerequisite Good data, know how to apply method,
- Often employed to adjust for or overcome data issues: selectivity, missing data.
- Overcome a roadblock
- SUCCESS STORY: Value energy star program
- Monetary value of program to members?
- Issues: self selection, intangible value (does market place value above and beyond measures take to save money?)
- Approach: statistical model accounted fro self selection using theoretical measure of intangible value (Tobins Q)

- Evaluation was successful
- Impact of enforcement on water quality
- issues: complex path to the outcome, 2 way relationship
- Approach: 2 stage statistical model followed by use of a water quality engineering model
- Significant impact were found : lots of questions from program managers(recreate analysis with different methods)
- PAINFUL STORY(FAILED PROGRAM)
- Program collected data before and after program. Some selectivity in data
- Used PSM to estimate program effects program manager agreed on the method
- Found small impacts
- Lots of pich back from program(focused on method use)
- #2
- Program needed 2 things # to report to OMB under GPRA in near term
- Valid method for the use over the longer term
- Near term method: used on member self assessments
- Longer them method account for missing data and selectivity (meant to be valid approach)
- Problem: near term method found bigger impact
- WHAT HAPPENED
- Energy star, good education lots of time spent on educating the program managers , precedence
- Enforcement and Water Quality, peer review, willing to explain approach, for re-analysis, precedence
- Agreement on method but didn't have acceptance, didn't educate well enough
- 2ND FAILED PROGRAM didn't educate well enough and peer review was done way after program.
- LESSONS
- Don't rely on "wow" factor. Program managers may or may not be impressed with the method and don't really care about method unless the results show the program is ineffective
- When using complex method, the methods is always under scrutiny and the method is never in the background, when it should be
- Agreement isn't the same as acceptance (must buy in on method)
- Within project peer review is valuable. Get reviewer as close to the program as possible.
- BEST PRACTICE
- Cross validate, show precedence, push for use of peer review, develop plain English descriptions { translate method into English, help manager understand that the methods is the most appropriate technique}
- ROLE AS EVALUATOR:
- Object analyst, salesperson (able to explain), punching bag (able to put why to a method)
- VALUE ADDED?

- Objectivity, appropriate method (apply method that will proved a valid answer to the question), what should be the value added? Education on method (be a salesperson!) and get buy in on method up front (agreement plus acceptance)

QUESTIONS

1. How to stop a program that has a fishing expedition? (knows we are looking for methods and not results) The intermediary division will stop a fishing expedition, if it happens. There are ways to avoid if while working with evaluation supports division.
2. If qualitative and quantitative methods are very different and one shows less good results do u ignore (delete) the results of the analysis with poor results? No
3. Education is important. Critical to involve customer in counterfactual area, but what do you do if a customer does not want to be evaluated? Must explain to them the value of evaluation. And show that this evaluation will help them to generate recommendations to their program. Hard to sell, but customer may be reluctant but knows it has to be done.
4. Does stakeholder involvement hurt or help the program manager by leading it down the wrong path? When involved, the program manager does want to try and answer the evaluation questions. Broader questions usually aren't asked, but why? Its important. Stakeholders don't ask the more important questions.