

14.11: Experiments in Political Science

Prof. Esther Duflo

May 9, 2006

Voting is a paradoxical behavior: the chance of being the pivotal voter in an election is close to zero, and yet people do vote.... So this is a place where “marketing” is expected to play a key role.

There is increasing interest in what gets people to vote.

The methods:

-canvassing

-phone

-mass mailing

The content of the message:

-reminder

-duty

-interest: either reminding that the election MAY be pivotal, or presenting other reasons why one should vote.

1 Voter Mobilization

Big names in this game: Don Green and Alan Gerber at Yale. Recent book: “Get Out the Vote.” Their analysis of the experiments just before the 2000 election was incorporated in both the 2000 campaign and the 2004 campaign.

1.1 The effect of canvassing, telephone calls, and direct mail on voter turnout: A field experiment

Sample: 29,380 individuals who are registered to vote.

Experimental design: See figure – A rather intricate 2*2*4 design to answer many questions. One note: Those who received more mailings also received more phone calls if they were assigned to the phone calls group.

On top of this, the mailings, canvassing, and phone call procedures varied according to the message that was provided:

-Civic duty

-Close election

-Politicians may ignore your neighborhood if people in this neighborhood do not vote

(Subjects who received one type of appeal in one form received the same type of appeal in the other forms.)

Consider this design. What are the questions we can answer? Which cells do we need to compare?

Strategy and results: Canvassing

Problem the experiment is facing: imperfect compliance.

Table 3 has a very clear description of the design.

Where do we read the *real* results of the experiments. What is the name of this estimator?

How does it compare to the results from the top panel? The bottom panel?

Why is the effect in the top panel larger?

Table 4: effect of the different types of messages. Conclusion?

Extending the strategy to other approaches

If you remember, the ratio in the bottom of Table 3 is an instrumental variable estimate: the instrument is whether or not you were randomly assigned to being in the canvassed group, and the variable of interest is whether you were in fact reached?

$$Y = \alpha_0 + \alpha_1 X_1 + \epsilon$$

Instrument for X_1 (canvassed) is T_1 (intended to canvass).

Now we have many treatments, so we may want to look at all of them in one regression:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_4 + \epsilon$$

where the various “intention to treats” are used as instruments for the various treatments (we can now also include control variables).

Notes:

-in this regression, someone who is sampled to be canvassed and mailed has a 1 for both of these treatments.

- X_2 and X_4 are continuous variables (0, 1, 2, 3) – what could we have done instead?

Comment on the results. What are the salient points?

If we want to know whether the latter matters more if people have been canvassed, what should we do? (We want to write down the equation and the instruments that will be useful.)

(As it turns out, the impact appears to be additive: no interaction effects were detected.)

1.2 Comparison between experimental and non-experimental estimates: The example of the impact of calls

Voting experiments offer the opportunity to compare experimental estimates to non-experimental estimates (something we maybe should have done in the first lecture!).

Non-experimental estimates of the effect of calls on voting

Many papers trying to look at the effect of voting campaigns compare people who were actually reached on the phone to people who were not reached.

What problems can this comparison face? What are the specific sources of bias we should worry about?

When you do not have an experiment, what can you do to try to make the “treated” and the “control” as comparable as possible?

A standard approach is to introduce control variables in a regression.

$$Y = \alpha_0 + \alpha_1 X_1 + W\beta\epsilon$$

What would you want to put in W ?

(In Table 3, we have the set of covariates that the authors are using.)

Alternatively, if we have a very large data set, we can try to form “cells” representing all the possible combinations of the values of W .

(i.e. compare the 20-year-old persons who live in households of two people, just registered, and live in California.)

In each cell, we will compare “treated” (those who were reached on the phone) and “control” (those who were not reached on the phone).

The identification assumption is that, within each cell, absent the call, the propensity to vote is the same for those who were reached as for those who were not reached. Is that reasonably plausible?

If we do that for each cell, we will get a very large number of estimates (one per cell). How should we put them together?

The solution is to take a weighted average of all the effect sizes estimated in each cell, weighted by the size of the cell (or the number of treatment people in a cell, to get the effect of the treatment on the treated).

What happens when there is a cell which has only controls and no treatment, or vice versa??

This method of estimation is called a “matching estimator” (because we match observations within cells).

How does it compare to our old-fashioned regression? When is it likely to differ a lot? If we are worried that even after controlling for all the W variables, there may still be unobservables that are different in treatment and control groups, does matching solve that problem?

Study design

The study was set up to compare experimental and matching estimates of the effect of calling people on whether or not they vote.

The experiment was conducted in Iowa and Michigan before the 2002 midterm election.

Two national phone banks were hired to read a non-partisan GOTV message to individuals in the treatment group. Individuals were randomly assigned to be called, but as before, not everyone could be reached. A sample of unlisted numbers were randomly added to the treatment and control group (they could not be reached so they should not matter for the experimental estimates, but they are there to study the impact they have on the non-experimental estimates).

Table 1: sample size and summary statistics

Results

First of all, how do we find out the experimental estimate of the effect of being called on voting?

Results are in Table 5: Conclusions? Does this fit with what we previously saw?

Now, the non-experimental estimates:

OLS in Table 6: Regress turnout on whether someone was actually contacted or not. Conclusions?

Matching in Table 7, and Table 9, including voting history. Conclusions?

Why is matching still different from the experimental estimate?

Look at Table 8: Why did they do this analysis and what clues does it provide?

1.3 Can asking people whether they are going to vote increase their propensity to vote?

Another way to increase the propensity to vote is to rely on the desire for:

-looking good in front of others.

-internal consistency (after promising to do something, you do not remember exactly why you promised, and you certainly do not want to remember having promised to do something just to look good in front of someone).

A class project in an experiment class studies the impact this has on the decision to register to vote and to actually vote.

First experiment: Voter registration. They call students and find out whether they are registered to vote or not. If they are not (66 students out of 419 were not registered), they assign them to two possible treatments:

-Ask them if they know where they can register, and if they don't know, tell them.

-For the treatment group: ask them whether they plan to register.

Results in Table 1: goes in the expected direction, but not significant (very small sample).

Second experiment: Voting. They call registered students and apply the same treatment and control. Effects are larger (at least in percentage point terms) and significant.

Nice experiment on this topic is described in the Cialdini book: Pallak, Cook and Sullivan, "Commitment and Energy Conservation."

The paper was published in 1975 (study was conducted in 1973, before you were born...), and I could not get ahold of it, so I do not even know whether it was randomized, etc.

But the idea is the following:

-group 1 is control

-group 2 was asked to make a commitment to reduce energy, and was assured anonymity.
-group 3 was asked to make a commitment to reduce energy, and they were told that their names would be published in the newspaper if they made a commitment.

Meter readings were collected by the local utility company and evaluated by Pallak, Cook and Sullivan. Consumption of electricity and natural gas for participating households was initially tracked for a one-month period. In order to determine whether changes in consumption levels would persist over time or whether they were only a result of “energy conscious” behavior for a few days after the home visit, the meter readings were tracked for a full 12-month period.

After one month, those in group 3 had reduced consumption of energy, those in group 2 had not.

At that point, those in group 3 received a letter saying that their names would not be published after all.

Their consumption REMAINED lower!

Interpretation?

By the way: nice theory paper by Benabou and Tirole, “Incentives and Pro-social Behavior” delves much deeper into models that take this desire for consistency, self-delusion, and desire to look good, into account (this ties to the potentially perverse effect of giving incentives to do good things that we have discussed earlier in the semester).

2 What appeals to the voter?

The voter mobilization experiments have a little bit to say about that.

They tried different messages:

- civic duty
- possibility of being pivotal
- “instrumental” voting

Not much evidence either way.

Paper “Clientelism and Voting Behavior” directly addresses the issues of whether people vote along these lines, in a unique experiment in Benin.

Democracy started in Benin in 1990, after 18 years of military dictatorship.

Experiment was conducted in 2001 during the first round of the presidential election, with the

two leading candidates (Kerekou–incumbent president, and Soglo–opposition), in their respective strongholds.

The objective of the paper is to measure whether electoral gains or losses associated with campaign promises are based on clientelism (transfer to village or region) as opposed to ones based on public policy (public goods).

8 non-competitive districts were selected for the experiment (4 dominated by each major candidate).

In each district, two “treatment villages” were chosen (it is not clear whether this was random or not) and in those village, they varied the platform: one clientelist platform, and one public good platform. The same team administered both.

Results:

-under most conditions, clientelists did better than control

-under most conditions (except south), public policy did worse than control (south result is explained because the south is dominated by a national opposition party, rather than a local party, and they are more credible when delivering a public good message).

-women are much more pro-public policy!