# Field Experiments

Peter Howley

Professor of Behavioral Economics

Leeds University Business School

# Recap – observational data

- Bivariate correlations and inferential statistics

  *- easy to do*

  *- gives an initial understanding of the relationship between variables*

  *- need to be aware of pitfalls surrounding correlation v causality*

    *- Omitted variable bias (third variable problem - confounding)*

    *- Bi-directional causality*

- Multivariate regression analysis

  *- big advantage is that it can help you 'control' for the influence of omitted variables*

  *- potential limitation (in certain circumstances) is that it can be hard to 'control' for everything*

  *- not very helpful when dealing with bi-directional causality*

# Field Experiments

By the end of this lecture, you should be able to:

*Describe the process behind an experimental design.*

*Understand how they work and how they can be used to determine 'causality'*

*Understand the meaning behind:* ***treatment*** *and* ***control group***, ***random selection***, ***selection bias***

*Get a 'feel' for how field experiments are used in the social sciences (some fascinating applications!)*

# Treatment v control group

- Field experiments represents the combination of two methodological strategies – **experimental designs** and **fieldwork**

**Experimental designs:** Individuals are randomly assigned *(by the researcher)* into a *treatment* and a *control* group

**Treatment group or groups**: The group receiving a treatment of some description, e.g. new drug, participate in a new policy program etc. etc.

**Control group**: A **control** is a group to which an experimental treatment is **NOT** administered. It serves as a ***reference mark for comparison or counterfactual*** (e.g., a group of subjects that do not receive the "new" drug, or a group of subjects that is given a placebo).

# Treatment v control group

- The difference between the two sample means (**treatment v control group**) provides the causal effect of the 'treatment'

  - inferential statistics (remember these! - e.g. t-tests) can be simply used to determine if any observed differences between the treatment and control group are statistically significant

- **In essence, field experiments provide useful methods of constructing the proper counterfactual (control group).**

  - *e.g. what would have happened if individuals were not treated*

# Key features of an experimental design

# Key features of an experimental design

- **Random assignment**

# Key features of an experimental design

- **Random assignment**

- **Random assignment**

# Key features of an experimental design

- **Random assignment**
- **Random assignment**
- **Random assignment**

# Experimental designs

**Why is random assignment so important?**

- If properly done it should eliminate potential for omitted variable bias

  - With **random assignment** you expect that both groups will have the same characteristics, e.g. level of motivation, ability or attitudes towards health, health related behaviors etc.

- Previously the view among economists was that statistical methods could be used to overcome the problems with observational data - e.g. regression adjustment (as we saw last week) to control for omitted variables, but field experiments offer a more powerful and reliable approach when they can be used

# Lets look at some applications!

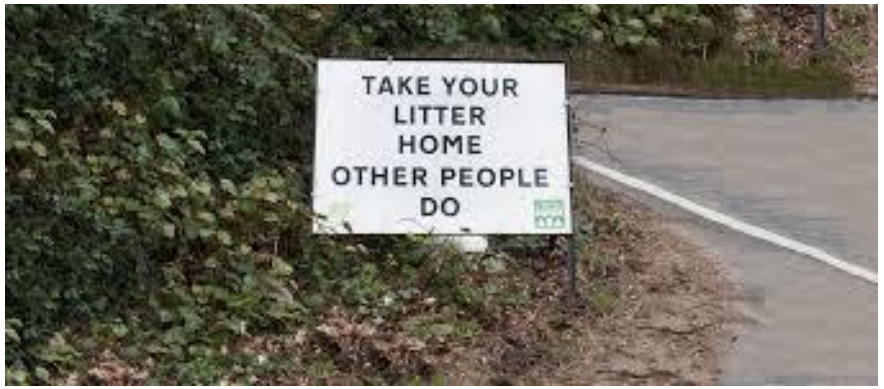
- Many field experiments test the effect of simply manipulating the information that people receive (often labeled as *informational treatments*) to see how that affects their subsequent behavior

- Let's look at an example in the area of environmental conservation

# Electricity Consumption

- Randomly chosen households were sent a letter every 1, 2, or 3 months with
  - info on electricity use compared to average household with similar characteristics (same area, same house size, etc.) and to "efficient" household with similar characteristics
  - classification "great" + two smiley faces, "good" with one smiley face, or "below average" with no smiley face
  - tips for reducing energy use

- Control group – just normal letter without comparison to neighbours




- Alcott (2011): Available here: https://www.sciencedirect.com/science/article/pii/S0047272711000478

# Electricity Consumption

- Letter reduces consumption for treatment relative to control group by ≈2%

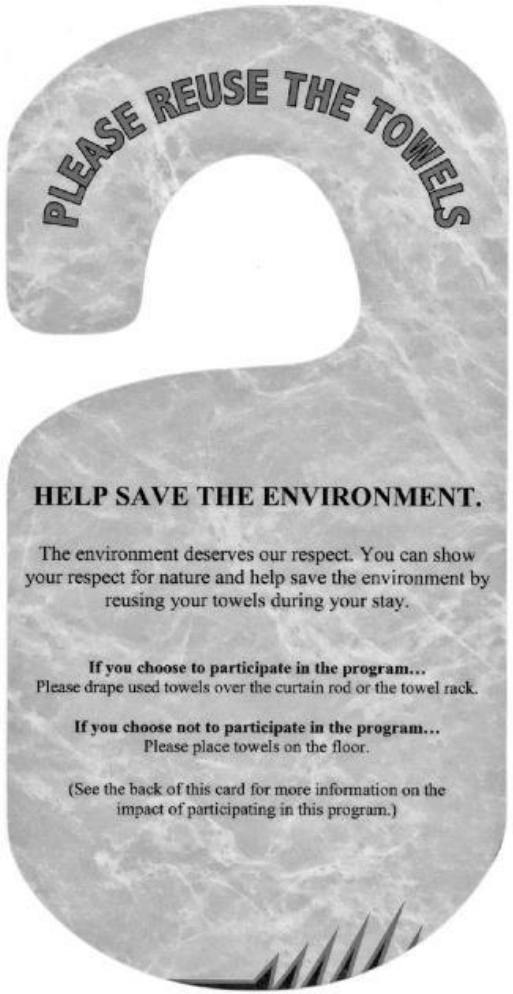
- Effect is larger for households with higher initial consumption

  *≈0% for households with lowest initial consumption (fortunately, they do not increase consumption after learning it's below average!)*

  *≈6% for households with highest initial consumption*

- Effect is persistent, especially if letters continue to be sent

- Can you think of any hypothesis to explain why these letters changed behavior? – rational actor agent model would suggest that it would not.




Goldstein et al. (2008) examined the effectiveness of signs requesting hotel guest' participation in an environmental conservation program

Subjects ***randomly assigned*** and given one of two potential messages

1) Industry standard (see picture): Focuses solely on benefits of reusing towels for environmental protection – **HELP SAVE THE ENVIRONMENT**
2) Normative appeal: **JOIN YOUR FELLOW GUESTS IN HELPING TO SAVE THE ENVIRONMENT**

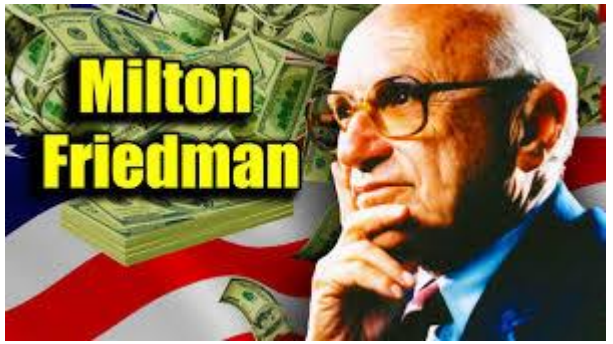
**Results**: Descriptive norm condition yielded a significantly higher towel reuse rate (44%) than the environmental protection condition

Goldstein et al. (2008): https://assets.csom.umn.edu/assets/118359.pdf

# Social Policy

- Negative income tax

- first large scale social science experiment ever conducted




- Individuals with incomes lower than the minimum would have their incomes supplemented by the tax system (a negative income tax).

    - in effect guarantees everyone a minimum income

- Would the guaranteed income lead to **mass emigration** from the labor force to public assistance, as low-wage workers lost the incentive to work?

# Negative income tax

- In four separate NIT experiments, over eight thousand households were *randomly assigned* to receive guaranteed incomes and others to a control group.

  -collected income/word data from **treatment and control groups** to see if NIT affected how people behaved

- The NIT experiment showed that a guaranteed minimum income had weak to moderate effects on labor effort.

  *- significantly less effect on behavior than would be predicted by standard neo-classical economics?*

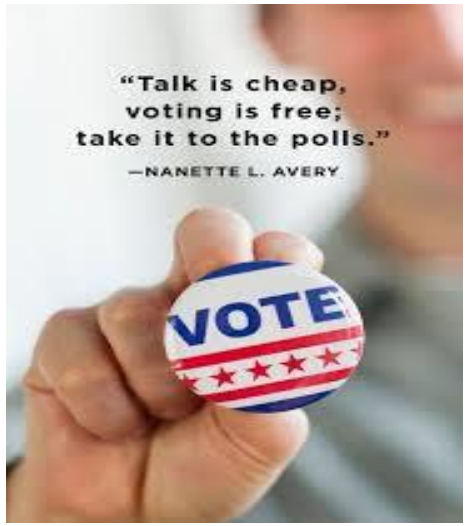
- Overall a 13% reduction in work effort– much of it in terms of less hours than leaving the labour force altogether

# Field Experiment on Voter Turnout

- Aim is to examine to what extent 'canvassing' by *personal contact*, *mail* and *telephone* can affect voter turnout
  - Researchers hypothesised that canvassing by personal contact will have a bigger impact on voter turnout than canvassing by direct mail or telephone
  - unclear how much difference there would be

- Previous work compared differences between the voting rates of those canvassed versus those who weren't and found that those canvassed had higher voting rates

  - what's the problem here?




Gerber and Green, (2000): https://isps.yale.edu/sites/default/files/publication/2012/12/ISPS00-001.pdf

# Field Experiment on Voter Turnout

- Aim is to examine to what extent 'canvassing' by *personal contact*, *mail* and *telephone* can affect voter turnout
  - Researchers hypothesised that canvassing by personal contact will have a bigger impact on voter turnout than canvassing by direct mail or telephone
  - unclear how much difference there would be

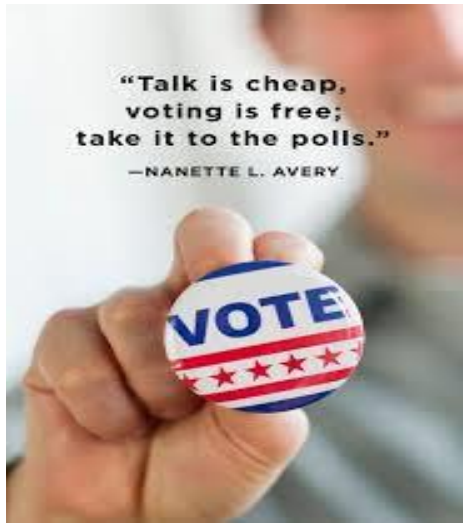
- Previous work compared differences between the voting rates of those canvassed versus those who weren't and found that those canvassed had higher voting rates

  - what's the problem here?

Multivariate regression analysis could be used to hold other factors constant

Still no guarantee (indeed very unlikely that you could in this case)

that you can control for everything




Gerber and Green, (2000): https://isps.yale.edu/sites/default/files/publication/2012/12/ISPS00-001.pdf

# Field Experiment on Voter Turnout

Collected a complete list of registered voters in New Haven, Connecticut

- divided into a series of treatment and control (not contacted) groups


- Test effect of encouraging people to vote by
  - ***canvassing (appearing at a person's house),***
  - ***telephone call***
  - ***direct mail (up to three)***

Face to face contact found to increase voter turnout by 9.8 percentage points

Direct mail by 0.6 percentage points

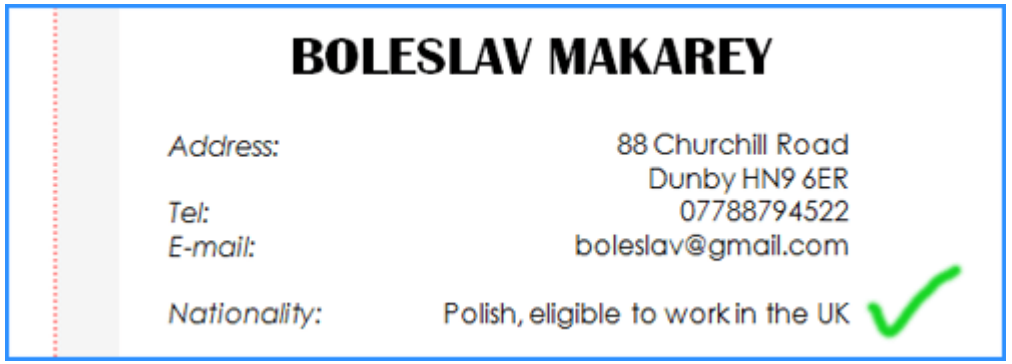
Telephone not found to have a statistically significant impact

# Discrimination

- https://www.youtube.com/watch?time_continue=5&v=ge7i60GuNRg
- https://www.youtube.com/watch?v=MuyMuLGXxTs

# Discrimination – more robust examination

- Researchers sent nearly 3,000 job applications under false identities in an attempt to discover if employers were discriminating against jobseekers with foreign names.




- Using names recognisably from three different communities – Nazia Mahmood, Mariam Namagembe and Alison Taylor – false identities were created with similar experience and qualifications. Every false applicant had British education and work histories.

  - people from ethnic minority groups were much less likely (21-29%) to be selected for interview

- A test of racial discrimination: http://natcen.ac.uk/media/20541/test-for-racial-discrimination.pdf

# Effectiveness of Financial and non-financial rewards

- Evaluate financial and non-financial rewards for HIV prevention and selling condoms.
- Field experiment randomized across Lusaka, Zambia.
- Predictions from mainstream neo-classical model of human behaviour ('rational man') would be that non-monetary rewards will not have much an impact here, but…
- No margin, no mission (Ashraf et al., 2014): https://www.sciencedirect.com/science/article/pii/S0047272714001546

# Financial and non-financial rewards

- In a field experiment in Zambia, hairstylists and barbers recruited by a public health organization to sell female condoms in their shops were randomly assigned to one of four groups receiving different awards based on condom sales

- ***Control group – no rewards***
- ***Large financial reward treatment***
- ***Small financial reward treatment***
- ***<u>Non financial reward (star) treatment</u>***

- What happened?

# Financial and non-financial rewards

After one year, hairdressers in the star treatment had sold **twice** as many condoms as hairdressers in any other group, on average.

- **For this group of individuals, the marginal utility of public recognition was higher than the marginal utility of money!**

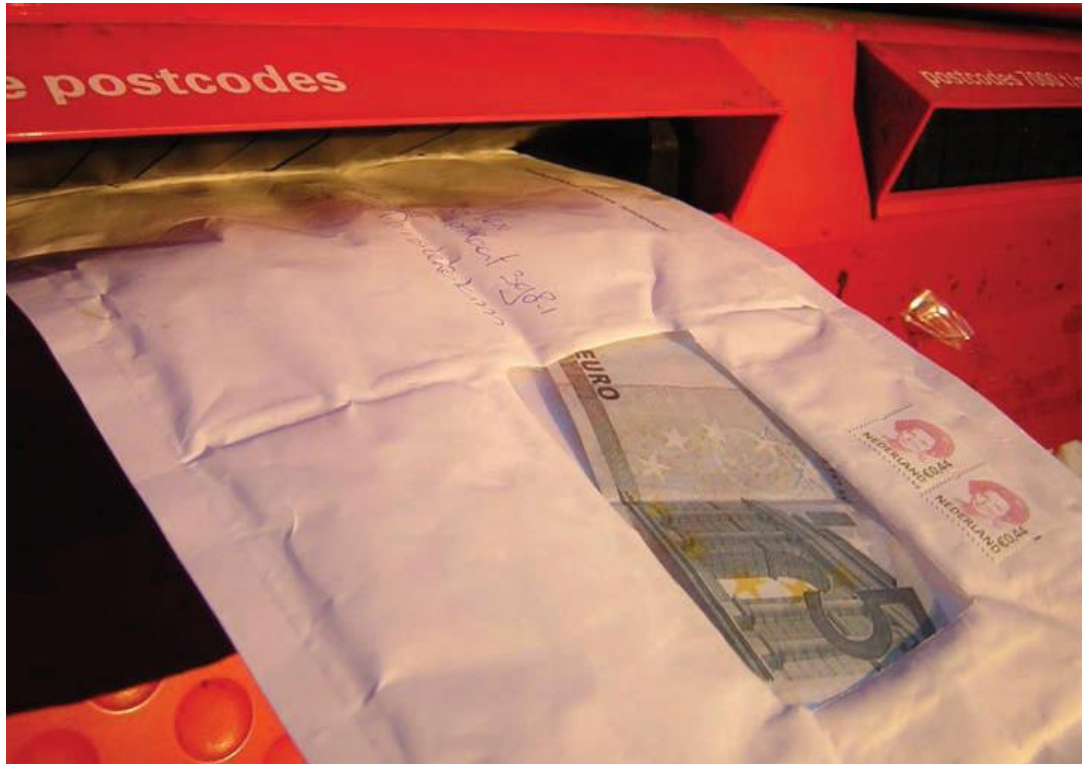
  **- how can you reconcile this with standard neoclassical theory**


## Spreading of disorder

Research question: To test the *broken windows theory* - hypothesis that signs of disorderly and petty criminal behavior trigger more disorderly and petty criminal behaviour

Keizer et al. (2008): Field experiments with two conditions (A - figure 1) and (B - figure 2).

- A leaflet was attached to the handlebar of bicycles
- Observed littering was much greater under B (69 v 33%) – why?
- The spreading of disorder (Keizer et al): https://www.rug.nl/staff/e.m.steg/keizerlindenbergstegdisorder.pdf




- **Bassline order condition** in which the mailbox was not covered with graffiti and the ground around the mailbox was clear

- **Two disorder conditions**

1. Mailbox covered with graffiti without litter on the ground
2. Mailbox without graffiti but litter on the ground

In baseline condition 13% stole the envelope, 27% stole in disorder condition 1 and 25% in disorder condition 2

Overall conclusion by the authors is that as certain norm-violating behaviours become more common, it will negatively influence conformity to other norms and rules

**The mere presence of graffiti more than doubled the number of people stealing!**

# Revision

- You will notice that I have left links to many of the studies that I discussed in this lecture. Pick a few of your favourites and read through.

- Ask yourself the following questions:

  *What was the key aim*

  *Why can it be classified as a field experiment*

  *How were people assigned to treatment and control groups and what was the 'treatment'*

  *Why was random assignment so important in that particular study*

  *What if any were the remaining threats to causal identification in these studies*

  *What did you think was the best thing about the study*

  *Can you think of any field experiments that you would like to try out*