SAVE Close Down Event

Trial evaluation

June 2019

University of Southampton

Scottish & Southern Electricity Networks

Trial evaluation

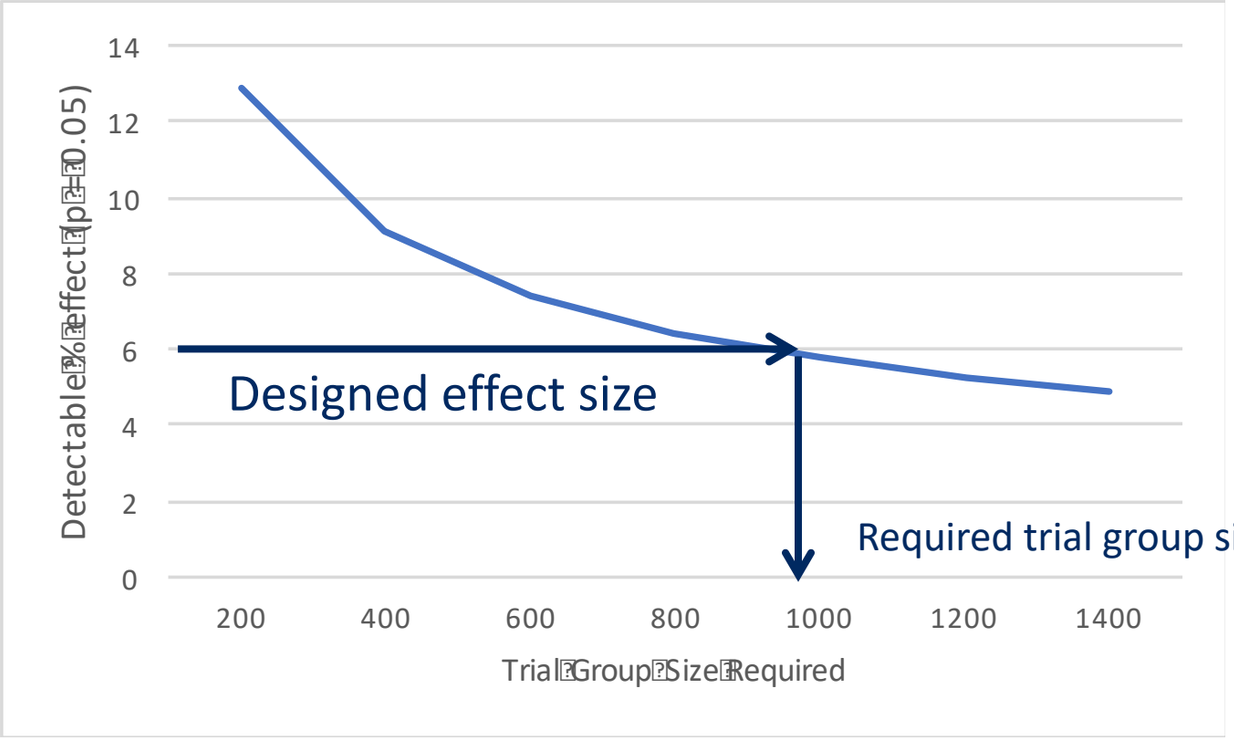
- Experimental design
  - SAVE: best practice trial design
  - Power analysis and sample size
  - Recruitment outcomes
- Trial evaluation challenges
  - Initial and revised analysis methods
  - Timescales and reference points
  - Attrition
- Summary and recommendations

What is ‘best practice’?

Table 1: Assessing behavioural interventions: A best practice framework (after Fredericks et al (2016))

Feature	Recommendation
Formulate Hypotheses	Clearly specify the expected effect of the interventions on behavior including their magnitude, direction and nature.
Program Design	Plan a sample size sufficient to give the statistical power required to test the hypotheses (to enable robust conclusions); Draw a random & representative sample (to enable generalisation) of the population of interest without self-selection (to avoid bias); Use a randomized control trial design wherever possible in order to be able to compare intervention with non-intervention groups; Randomly allocate participants to control or trial groups without self-selection (to avoid bias);
Methodology	Define and assess sample ‘representativeness’; Collect baseline data on key socio-economic and demographic attributes to assess sample ‘representativeness’; Establish that control and intervention groups are equivalent in key respects prior to interventions.

Statistical power and sample size



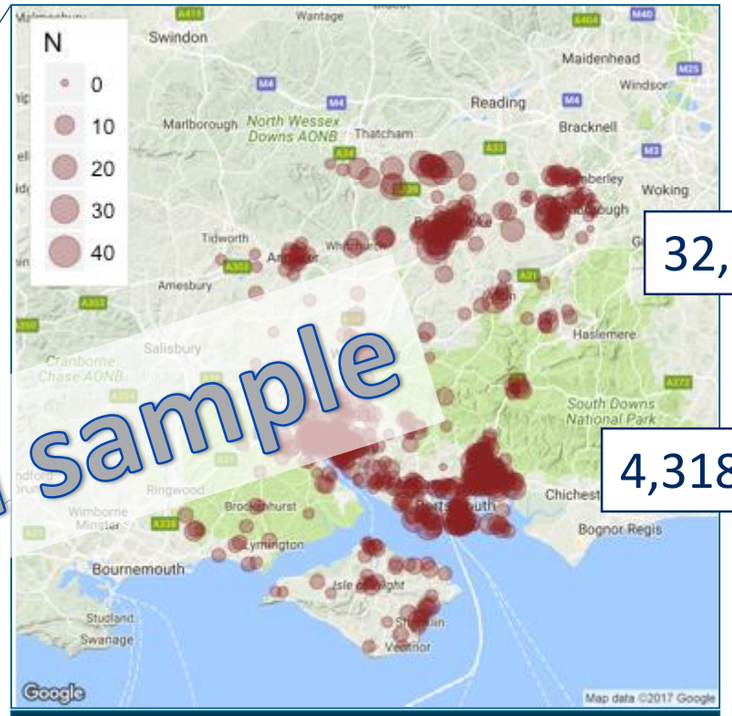
⇒ **Each** trial group > 1000
⇒ Control + 3 trial groups
⇒ Total sample > 4,000 households

Source: UoS analysis of Irish CER Domestic Demand Response pre-trial consumption data
Mean kWh 16:00 – 20:00 (“Evening peak”)
p = 0.05, P = 0.8

Sampling

- Hampshire, Isle of Wight, Southampton, Portsmouth
- Sampling stratified by the random selection of Census OAs within deprivation quintiles
- Random selection of 50 addresses from each OA
- Random allocation to treatment groups

=> Random sample



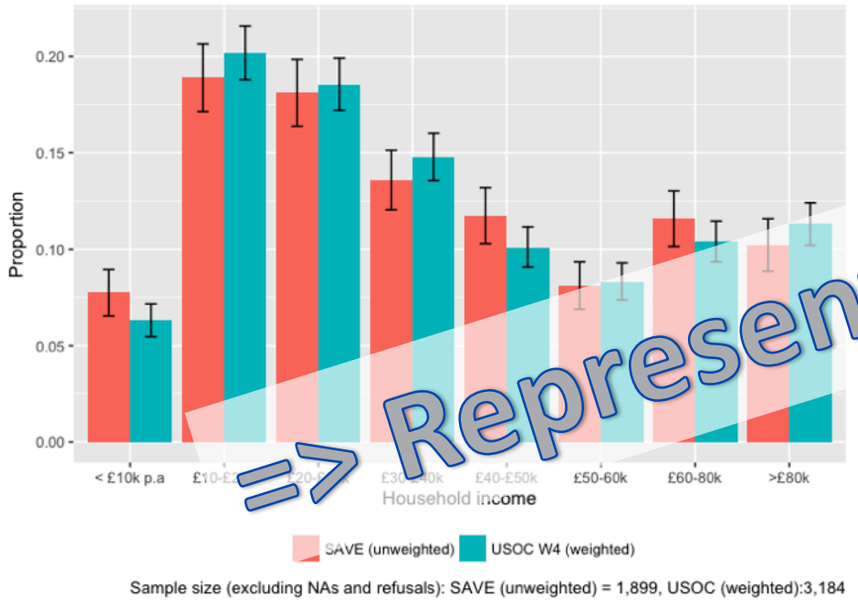
Geographical location of SAVE project trial participants

32,000 letters

4,318 households

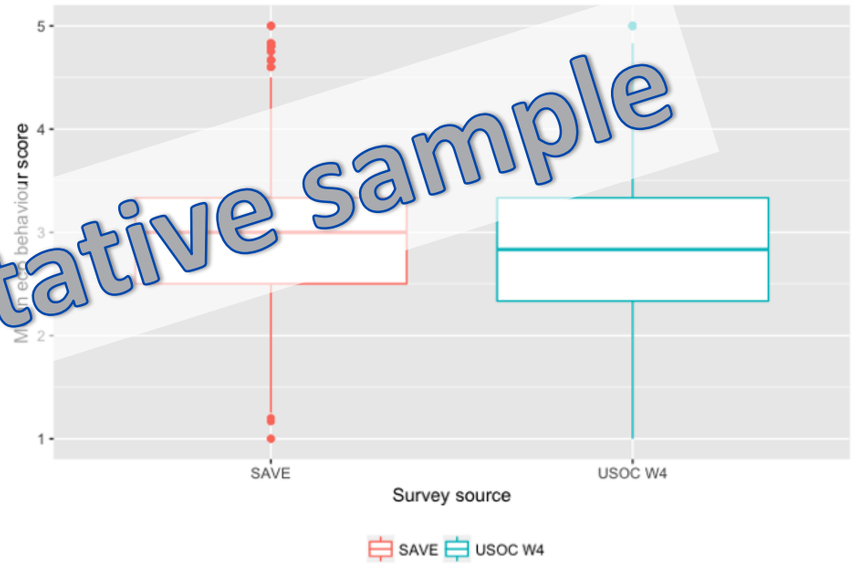
Recruitment outcomes: representative?

- Income



Sample size (excluding NAs and refusals): SAVE (unweighted) = 1,899, USOC (weighted):3,184

- Environmental attitudes



Sample size (excluding NAs): SAVE = 2,899, USOC:2,903

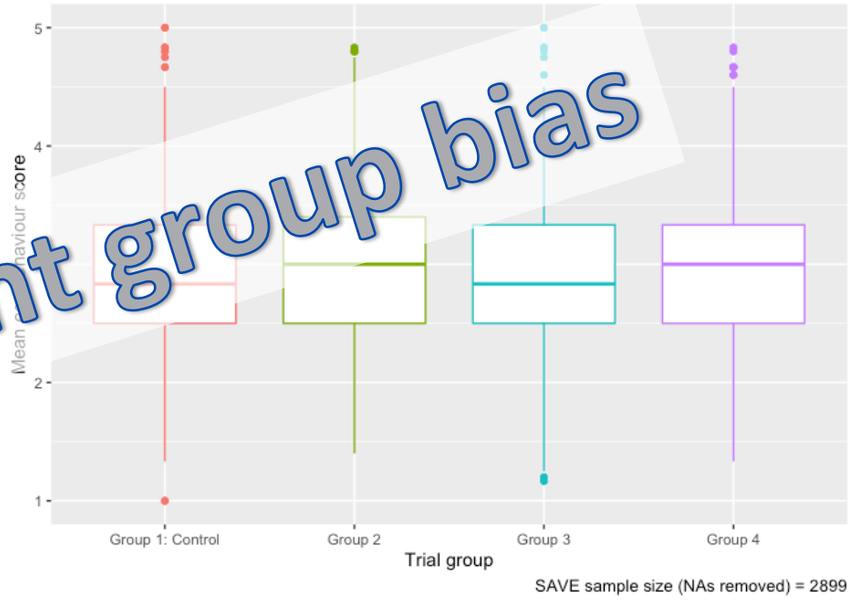
=> Representative sample

Error bars: 95% Confidence Intervals
Source: UoS analysis of SAVE vs Understanding Society Wave 4 sample for South East England (weighted for non-response)

Recruitment outcomes: biased?

- Electricity consumption

- Environmental attitudes



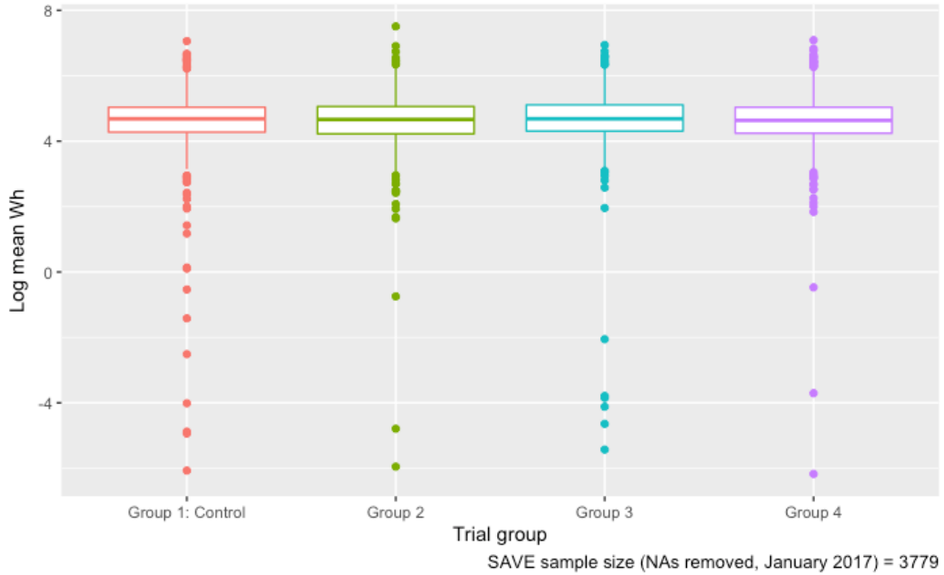
=> No treatment group bias

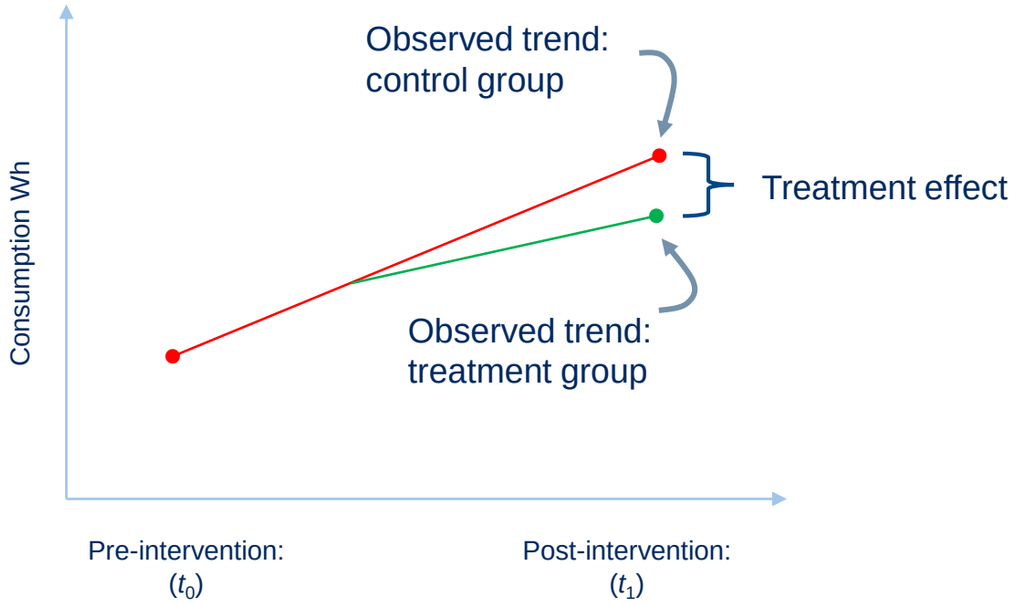
Error bars: 95% Confidence Intervals
Source: UoS analysis of SAVE vs Understanding Society Wave 4 sample for South East England (weighted for non-response)

Recruitment outcomes

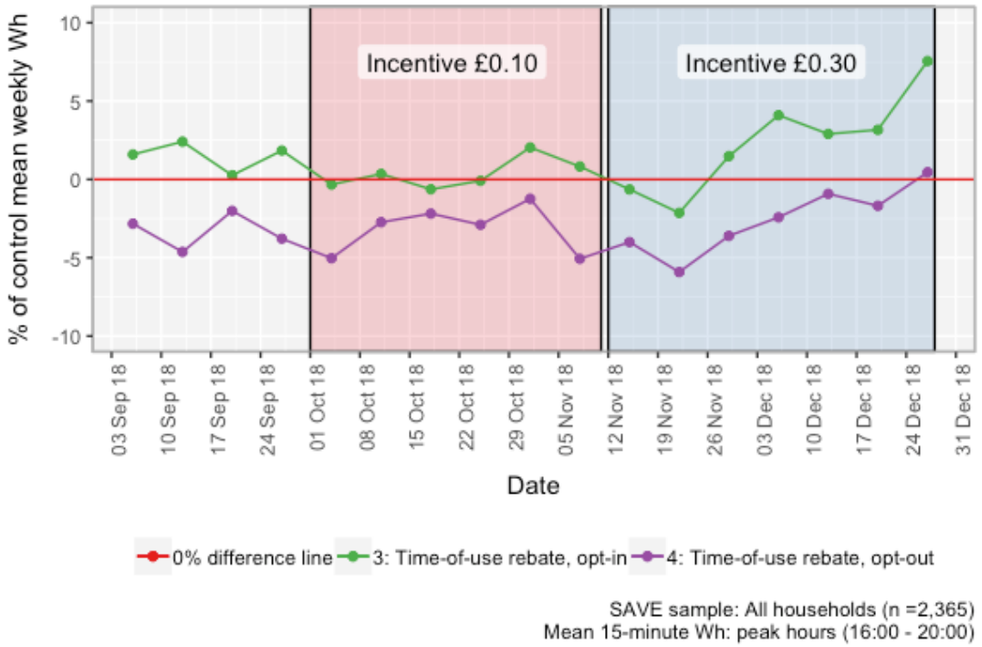
Large sample size	Statistically robust
Random allocation to trial groups	Equivalent groups: differences in consumption can be attributed to intervention
Random, representative sample	Results are generalisable to the wider population

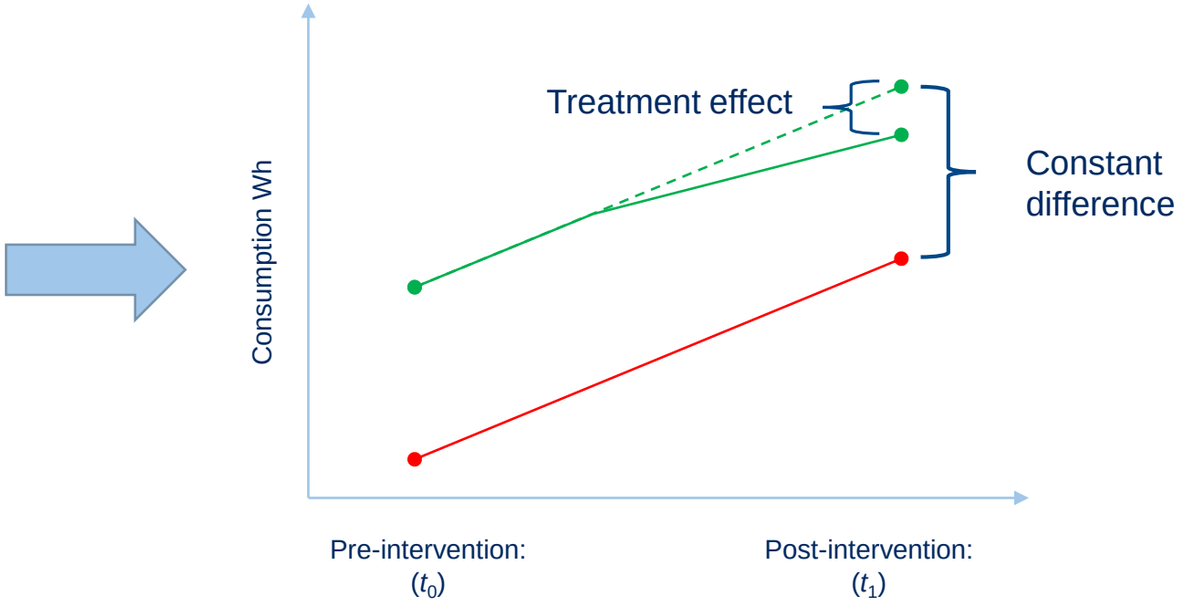
Analysis method – equivalent trial groups



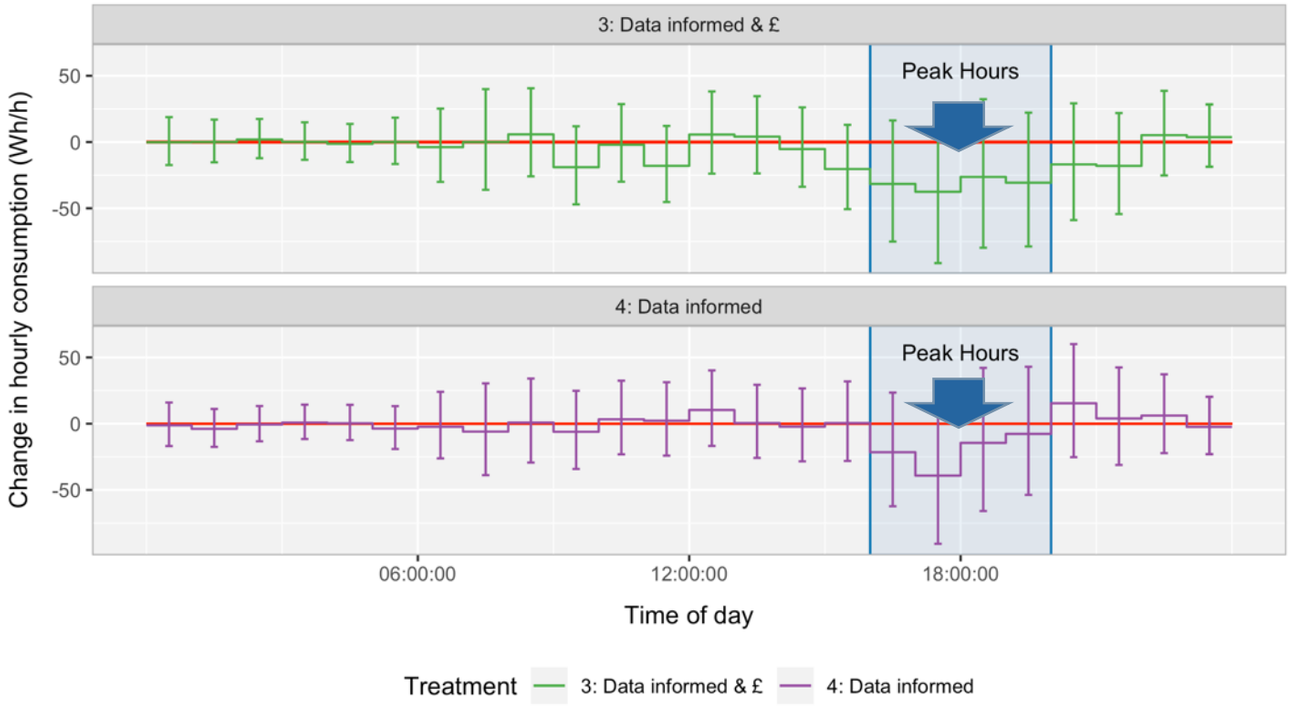


Analysis method – asymmetrical groups

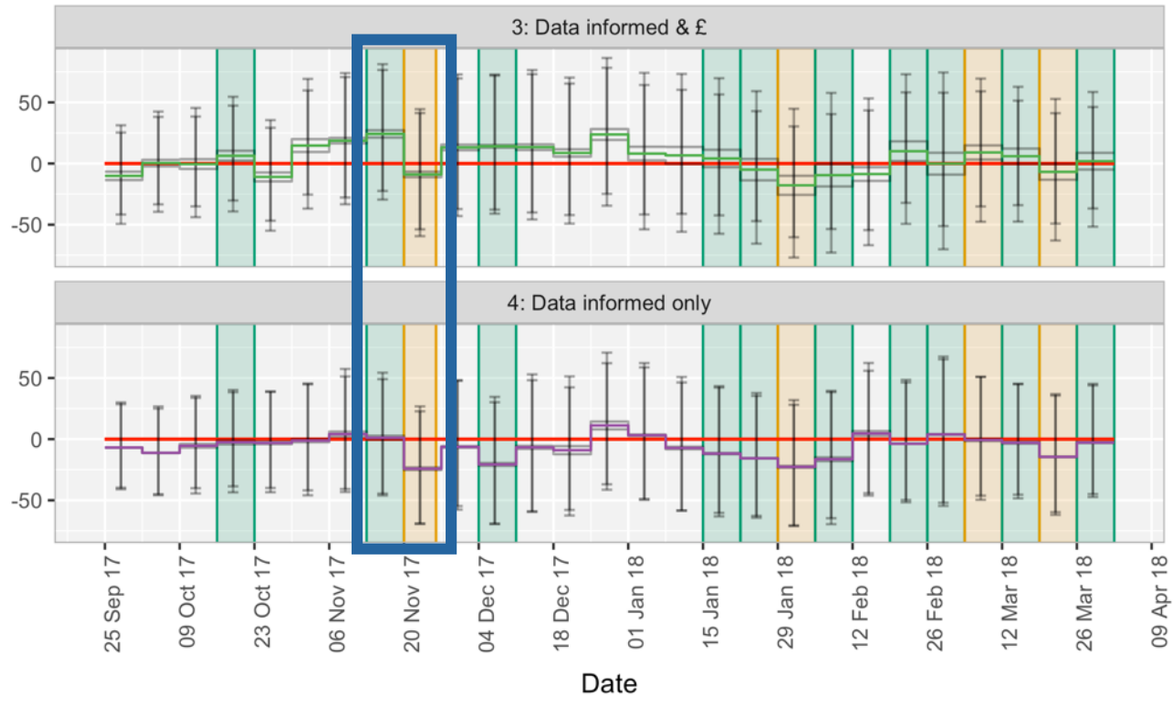




Timescales – short and long-term effects

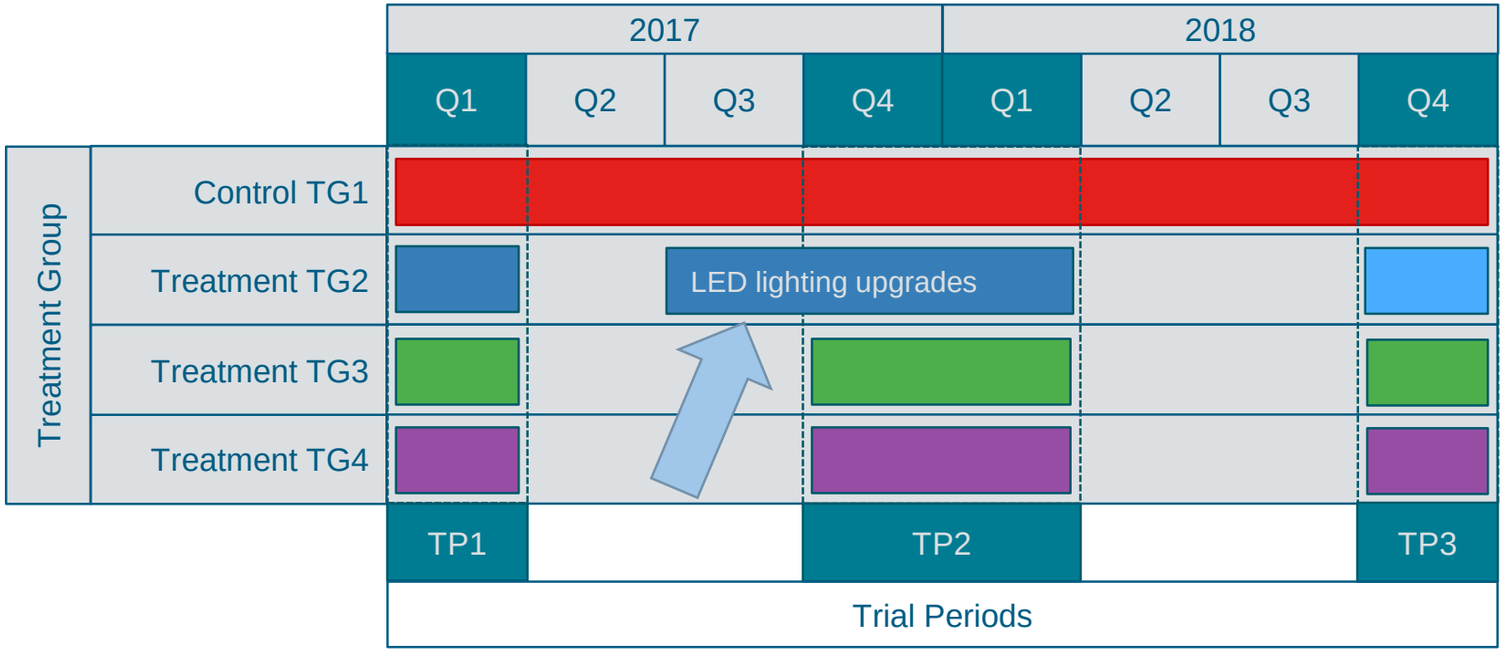


SAVE sample households: 2017-11-13 to 2017-11-26
Sample size: Control = 861, Treatment = 794 & 791
Error bars indicate 90 percent confidence interval for estimates

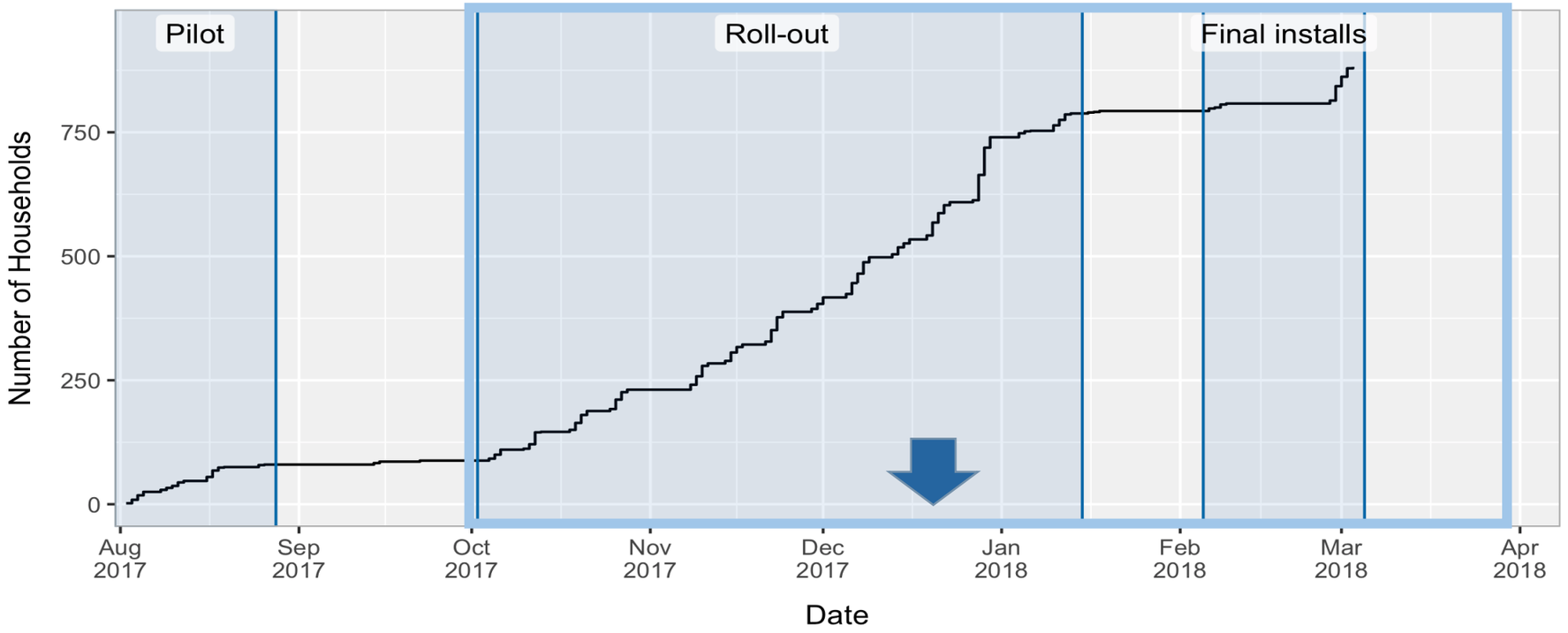


Error bars: 90% confidence interval for the estimates

Sample attrition

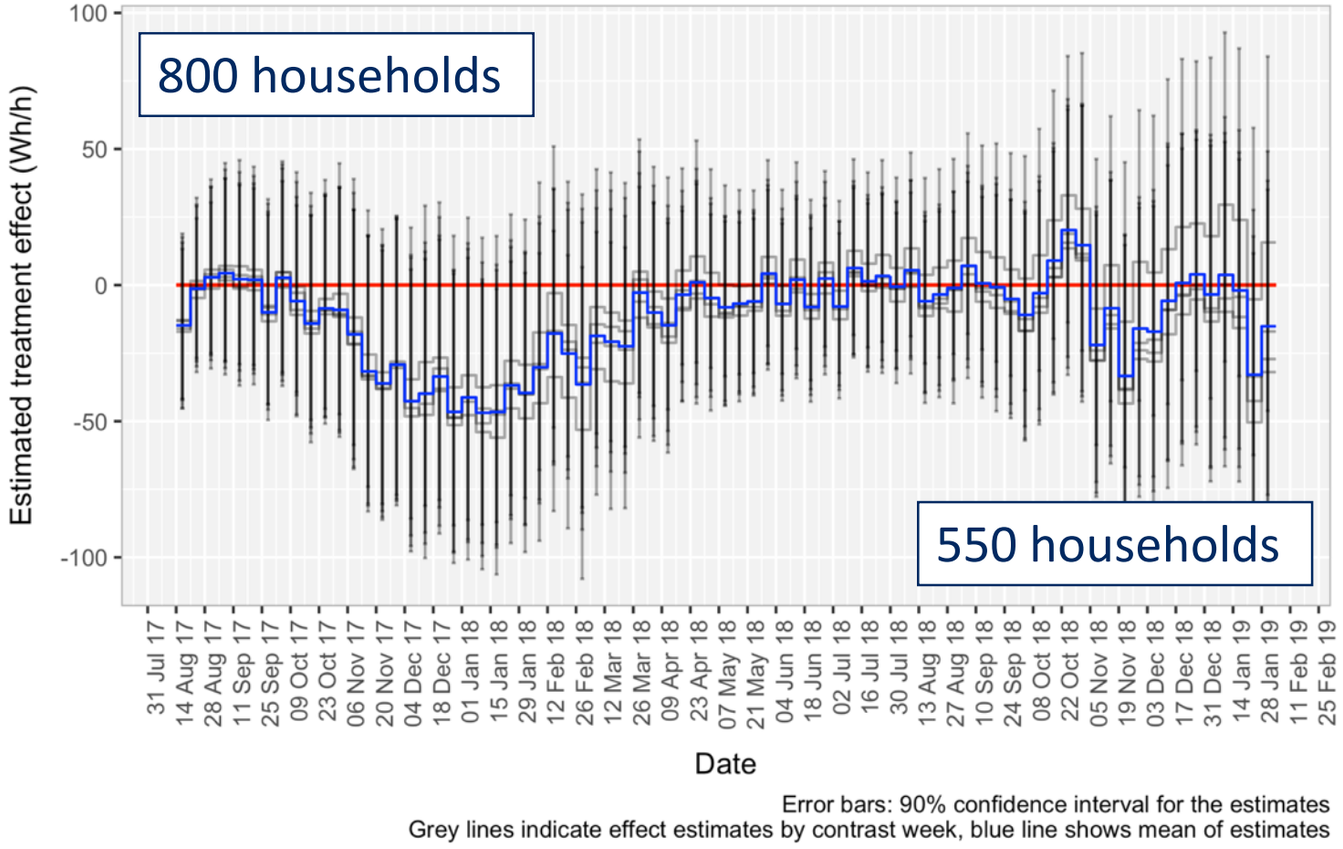


Sample attrition



Cumulative total of LED lightbulb installations
Shaded area denotes Trial Period 2

Sample attrition



Extended evaluation period, however this resulted in:

- Smaller sample
- Increased uncertainty in estimated treatment effects
- Difficulty in evaluating the maximum savings

Summary and recommendations

SAVE delivered a robust, best practice trial design to provide industry-leading evidence base for estimating and modelling demand response

- Even with careful design and implementation, the project faced evaluation challenges:
  - Small asymmetries between groups required a new analytical approach
  - Understanding responses to interventions required analysis across multiple time scales
  - Attrition and communications issues over the trial increased uncertainty
- Recommendations:
  - Plan for asymmetry in trial groups even for RCTs with equivalent trial groups at trial start
  - Be realistic about timescales around recruitment and interventions
  - Adapt analysis approaches to each intervention
  - Sample size: plan for attrition and communication issues

Thank you for listening.

t.w.rushby@soton.ac.uk

@tom_rushby

#SAVEClosedown