Social Media-Based Field Experiments: Individual-level assignment

Florian Foos

Assistant Professor in Political Behaviour
Department of Government
London School of Economics & Political Science

f.foos@lse.ac.uk
@florianfoos

July 27, 2021

Notation

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

- $d_i = 1$

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

- $d_i = 1$ means the ith subject receives the treatment.

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

- $d_i = 1$ means the ith subject receives the treatment.

- $d_i = 0$

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

- $d_i = 1$ means the ith subject receives the treatment.
- $d_i = 0$ means the ith subject does not receive the treatment.

Notation

- Indexing experimental subjects/units: subscript i refers to subjects 1 through N.

- Defining treatment: The variable d_i indicates whether the ith subject is treated.

- $d_i = 1$ means the ith subject receives the treatment.

- $d_i = 0$ means the ith subject does not receive the treatment.

- It is assumed that d_i is observed for every subject.

Recap: Potential outcomes

Recap: Potential outcomes

- $Y_i(1)$

Recap: Potential outcomes

- $Y_i(1)$ is the potential outcome if the ith subject was treated.

Recap: Potential outcomes

- $Y_i(1)$ is the potential outcome if the ith subject was treated.

- $Y_i(0)$

Recap: Potential outcomes

- $Y_i(1)$ is the potential outcome if the ith subject was treated.

- $Y_i(0)$ is the potential outcome if the ith subject was not treated.

Recap: Conditional potential outcomes

Recap: Conditional potential outcomes

- $Y_i(0) \mid d_i = 1$

Recap: Conditional potential outcomes

- $Y_i(0) \mid d_i = 1$: untreated potential outcome for subjects that receive the treatment.

Recap: Conditional potential outcomes

- $Y_i(0) \mid d_i = 1$: untreated potential outcome for subjects that receive the treatment.

- $Y_i(1) \mid D_i = 0$

Recap: Conditional potential outcomes

- $Y_i(0) \mid d_i = 1$: untreated potential outcome for subjects that receive the treatment.

- $Y_i(1) \mid D_i = 0$: treated potential outcome for subjects that would not receive the treatment under a hypothetical random assignment.

Random assignment

Random assignment

- "Experiment": a study in which subjects are assigned to treatment (and control) with a known probability between 0 and 1.

Random assignment

- "Experiment": a study in which subjects are assigned to treatment (and control) with a known probability between 0 and 1.

- Random assignment: the probability of assignment to treatment (and control) is equal for each subject.

Random assignment

- "Experiment": a study in which subjects are assigned to treatment (and control) with a known probability between 0 and 1.

- Random assignment: the probability of assignment to treatment (and control) is equal for each subject.

- That means no subjects has a higher probability to be treated than another subject. There is no attribute that can systematically predict treatment assignment.

Random Assignment

Random Assignment

- Types of random assignment:

Random Assignment

- Types of random assignment:
  1. Simple random assignment: Each subject is allocated to the treatment group with probability m/N, where m is the number of subjects assigned to treatment.

Random Assignment

- Types of random assignment:
  1. Simple random assignment: Each subject is allocated to the treatment group with probability m/N, where m is the number of subjects assigned to treatment.

  2. Complete random assignment (exactly m units are assigned to treatment).

Random Assignment

- Types of random assignment:

  1. Simple random assignment: Each subject is allocated to the treatment group with probability m/N, where m is the number of subjects assigned to treatment.

  2. Complete random assignment (exactly m units are assigned to treatment).

  3. Block-random assignment (exactly m units within each subgroup defined by pre-treatment covariate are assigned to treatment).
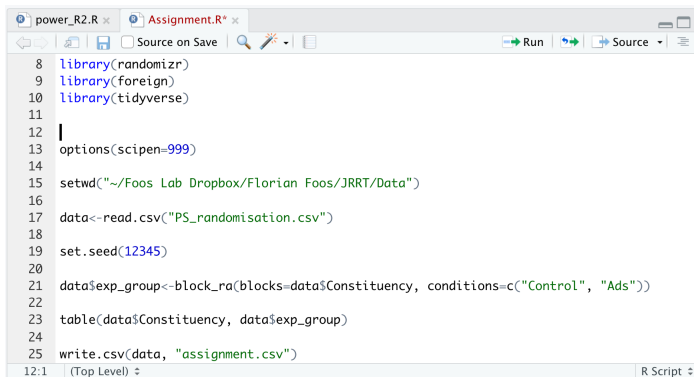
How to randomize?

How to randomize?

- Can we done in Excel using the RAND() function.

How to randomize?

- Can we done in Excel using the RAND() function.

- Much better: Use the *randomizr* package in R (Coppock, Cooper and Fultz 2019),

How to randomize?


```r
8   library(randomizr)
9   library(foreign)
10  library(tidyverse)
11
12
13  options(scipen=999)
14
15  setwd("~/Foos Lab Dropbox/Florian Foos/JRRT/Data")
16
17  data<-read.csv("PS_randomisation.csv")
18
19  set.seed(12345)
20
21  data$exp_group<-block_ra(blocks=data$Constituency, conditions=c("Control", "Ads"))
22
23  table(data$Constituency, data$exp_group)
24
25  write.csv(data, "assignment.csv")
```

How to randomize?

```
> table(data$Constituency, data$exp_group)
```

	Control	Ads
Aberdeen South	13	12
Alyn and Deeside	9	8
Angus	12	12
Ayr, Carrick and Cumnock	11	11
Banff and Buchan	14	14
Battersea	23	23
Bedford	13	14
Bermondsey and Old Southwark	14	15
Bolsover	12	11
Brentford and Isleworth	19	18
Carshalton and Wallington	12	12
Chipping Barnet	14	14
Clwyd South	10	9
Colne Valley	11	11
Croydon Central	10	10
Dagenham and Rainham	11	10
Darlington	7	7
Delyn	5	5
Don Valley	10	10
Eltham	11	11
Enfield North	10	10
Enfield Southgate	6	7
Erith and Thamesmead	7	7
Finchley and Golders Green	12	12
Great Grimsby	8	8
Harrow East	8	9
Hendon	9	8
High Peak	14	15
Ilford North	8	7
Keighley	12	12
Kensington	16	16
Moray	11	11
Ochil and South Perthshire	18	17
Putney	4	5

Some basics
Random assignment
Examples
Non-interference
Your own social media expriment

Munger 2017
Foos et al. 2020
Guess et al. 2021

Munger (2017)

Table 1 Experimental design and hypothesized effect sizes

	In-group	Out-group
Low followers	Medium effect	Small effect
High followers	Large effect	Medium effect

Some basics
Random assignment
Examples
Non-interference
Your own social media expriment

Munger 2017
Foos et al. 2020
Guess et al. 2021

Munger (2017)

Fig. 3 Treatments. **a** The treatment—black bot. **b** The bot applying the treatment—white bot

Foos et al. 2020

Some basics
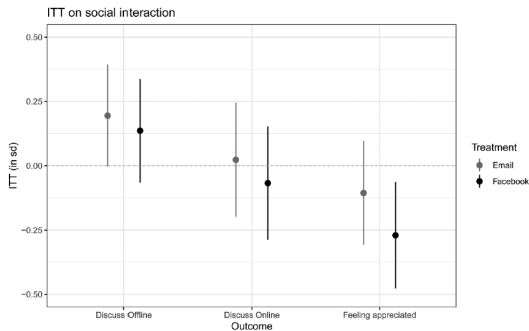Random assignment
Examples
Non-interference
Your own social media expriment

Munger 2017
Foos et al. 2020
Guess et al. 2021
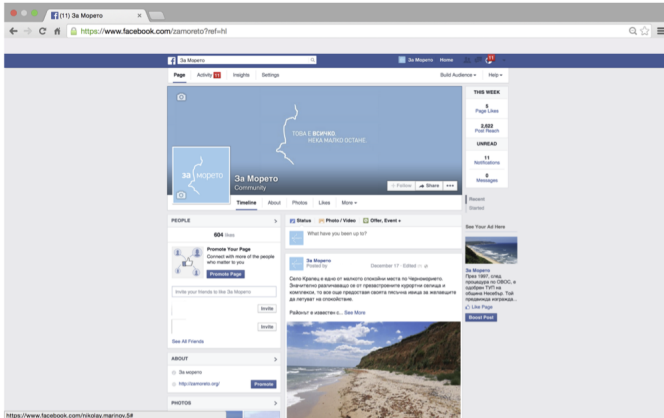
Foos et al. 2020

Figure 4. Effect of treatment assignment (ITTs) on social interaction—covariate-adjusted, 95 percent CIs.

Some basics
Random assignment
Examples
Non-interference
Your own social media expriment

Munger 2017
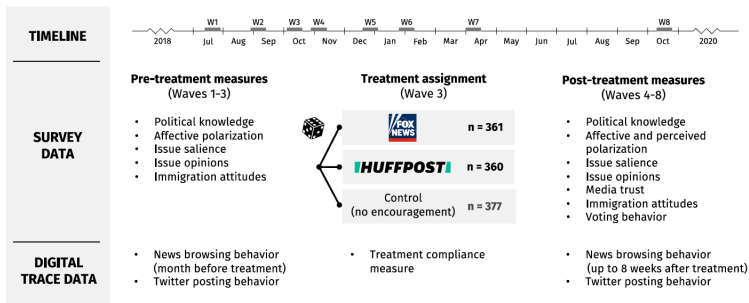Foos et al. 2020
Guess et al. 2021

Guess et al. 2021

Fig. 1. Overview of study design. Subjects in wave 3 who were randomly assigned to the Fox News or HuffPost encouragement groups were offered $8 in YouGov incentives to participate in the treatment.

Some basics
Random assignment
Examples
Non-interference
Your own social media expriment

Munger 2017
Foos et al. 2020
Guess et al. 2021

The three core assumptions

- Independence

- Excludability

- Non-interference

The three core assumptions

- Independence

- Excludability

- **Non-interference**

Non-interference assumption / SUTVA

- The values of the potential outcomes for subject i depend only on whether the subject itself is treated (whether d equals 1 or 0).

- Each subject is unaffected by the treatments and assignments of other units.

Non-interference assumption

Non-interference assumption

- $Y_i(d)$:

Non-interference assumption

- $Y_i(d)$: the potential outcome that would be expressed based on the treatment that subject i receives.

Non-interference assumption

- $Y_i(d)$: the potential outcome that would be expressed based on the treatment that subject i receives.

- $Y_i(\mathbf{d})$:

Non-interference assumption

- $Y_i(d)$: the potential outcome that would be expressed based on the treatment that subject i receives.

- $Y_i(\mathbf{d})$: the potential outcome that subject i would express based on the assignments that all other subjects receive.

Non-interference assumption

- $Y_i(d)$: the potential outcome that would be expressed based on the treatment that subject i receives.

- $Y_i(\mathbf{d})$: the potential outcome that subject i would express based on the assignments that all other subjects receive.

- $Y_i(d) = Y_i(\mathbf{d})$

Non-interference assumption

- $Y_i(d)$: the potential outcome that would be expressed based on the treatment that subject i receives.

- $Y_i(\mathbf{d})$: the potential outcome that subject i would express based on the assignments that all other subjects receive.

- $Y_i(d) = Y_i(\mathbf{d})$

- The equality means that subject i is unaffected by the treatment of other subjects.

Interference between subjects is common on social media

Interference between subjects is common on social media

- But what if there is interference/interaction between subjects (as is often the case on social media)?

Interference between subjects is common on social media

- But what if there is interference/interaction between subjects (as is often the case on social media)?

    1. Contagion

Interference between subjects is common on social media

- But what if there is interference/interaction between subjects (as is often the case on social media)?

    1. Contagion

    2. Communication

Interference between subjects is common on social media

- But what if there is interference/interaction between subjects (as is often the case on social media)?

  1. Contagion

  2. Communication

  3. Social comparison

Interference between subjects is common on social media

- But what if there is interference/interaction between subjects (as is often the case on social media)?

    1. Contagion

    2. Communication

    3. Social comparison

    4. Inference across platforms including displacement.

Interference between subjects

Interference between subjects

- What to do if potential outcomes are likely to depend on other subjects' treatment status?

Interference between subjects

- What to do if potential outcomes are likely to depend on other subjects' treatment status?

  1. We can expand the schedule of potential outcomes and redefine the estimator.

$Y_i(1, 1)$	$Y_i(0, 0)$
$Y_i(1, 0)$	$Y_i(0, 1)$

Interference between subjects

- What to do if potential outcomes are likely to depend on other subjects' treatment status?

  1. We can expand the schedule of potential outcomes and redefine the estimator.

$Y_i(1, 1)$	$Y_i(0, 0)$
$Y_i(1, 0)$	$Y_i(0, 1)$

  2. We can obtain a measure of the network pre-treatment and assign subjects in the same network together to treatment or control (cluster-random assignment)

If you could conduct a randomized experiment using social media, what would it be?

- What would the research question be?

- Which platform would you use?

- What would be the treatment?

- Any complications to keep in mind?

Time for questions.