POIR 613: Computational Social Science

Pablo Barberá

University of Southern California pablobarbera.com

Course website: pablobarbera.com/POIR613/

Today

- 1. Zoom discussion sessions (for now):
 - September 22nd
 - October 20th
 - November 10th
- 2. Experimental research in the digital age
- 3. Solutions for last week's challenge
- Webscraping

Experimental research in

the digital age

Experimental research in the digital age

Field experiments combine the control of laboratory experiments (high internal validity) with the generalizability of a real setting (external/convergent validity).

Challenge: cost, particularly if scale is sufficient to study high-variance social phenomena.

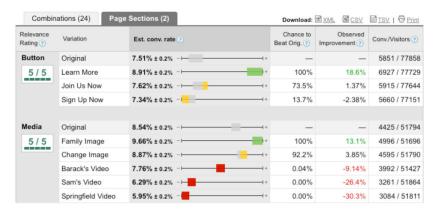
Digital technologies offer practical and cost-effective venues for conducting field experiments (aka A/B tests).

Given sufficient access and existence of software that allows randomization, researchers can study both short- and long-term effects of manipulations



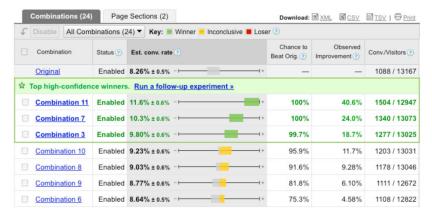


6 Media variation \times 4 button combinations = 24 combinations Which one do you think will get a higher conversion rate?



Outcome variable: sign-up rates

Dashboard shows sign-up rates for each separate variation



Dashboard shows sign-up rates for each separate variation

The winner



Original sign-up rate: 8.26% New sign-up rate: 11.6% Change: +40.6 lift in sign-up rate

10MM people signed-up through splash page during campaign

Without experiment, number would have been 7.2MM

That's 2.8MM fewer email addresses

Average donation per email address is \$21

 $2.8MM \times \$21 = \$60MM !!!$

Experimental research in the digital age

Experimental technologies for online interventions:

1. Email and text messages

- More likely to get subjects' attention
- e.g. Blair et al (2017): randomized text messages in India to encourage people to report corruption

2. Modified web interface

- Manipulation: platform features, exposure to information, display of specific web elements, etc.
- e.g. Bakshy et al (2012): social cues on FB ads

3. Bots

- Program or script that makes automated requests
- e.g. Munger (2016): reducing harassment on Twitter

4. Add-ons

- Additional software that nudges or tracks subjects
- e.g. Guess (2016): web tracking software to observe individuals' news consumption in response to monetary encouragement to seek information

Experimental research in the digital age

What can go wrong? (And potential solutions)

- Logging errors: covariate balance in pre-treatment variables, A/A tests
- 2. Novelty effects: longer experiments
- 3. Multiple testing: Bonferroni correction
- 4. High significance due to large sample sizes: Cohen's D
- SUTVA (interference between units): better research design
- 6. The 'free beer' problem: social science theory!

Side note: power calculations

- Power is the probability of detecting a specified effect size with specified sample characteristics (size and variability)
- Four interrelated components:
 - 1. Sample size
 - 2. Effect size you want to detect
 - 3. Desired significance level (false positive rate you're fine with)
 - 4. Power
- Before you run an experiment, you can compute necessary sample size assuming other 3 components:

```
> power.prop.test(p1=0.30, p2=0.35,
sig.level=0.05, power=0.80)
```