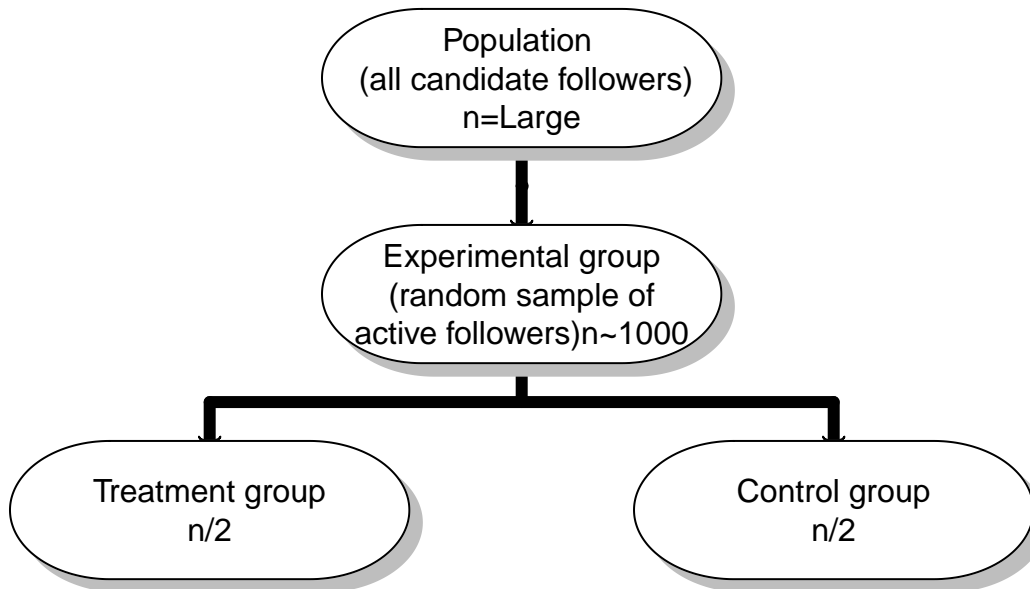# Proposal

Experimental Design

For [two] political candidates, we create a subset of their followers whose behaviour we observe. This subset is split into a treatment and control group as follows.

```
                    ┌─────────────────────────┐
                    │       Population         │
                    │ (all candidate followers)│
                    │        n=Large           │
                    └─────────────────────────┘
                                 │
                    ┌─────────────────────────┐
                    │   Experimental group     │
                    │   (random sample of      │
                    │ active followers)n~1000  │
                    └─────────────────────────┘
                      │                     │
          ┌───────────────────┐   ┌───────────────────┐
          │  Treatment group  │   │   Control group   │
          │        n/2        │   │        n/2        │
          └───────────────────┘   └───────────────────┘
```

Size of n

- The Rate limit for posting tweets is 2400 per day
- The Rate limit for direct messages is 1000 per day

We should identify a minimum experimental group size for statistical relevance and work from there.

Treatment Definition

We Provide information to the treatment group about the veracity of the claims made by the candidate they follow. Politifact.com won the Pulitzer prize for providing fact checking. They have an RSS feed for each candidate. Every time Politifact write a post about the candidate, we send a tweet summarising the fact-checking judgement and [providing a link] (this could make us look more like spam) to our treatment group.

Trump's last 10 checked statements go back to the $23/02/2016 = 0.83$ per day

Clinton's last 10 go back to $07/02/2016 = 0.35$ per day

Tweets are put into an "outbox" as statements are checked, and the outbox is sent to treatment groups at a steady rate.

Drawbacks

The Twitter Rules Issue

Guidelines

The general guidelines say this

> @Replies You can direct a Tweet at a specific Twitter user using @replies and mentions. The @reply feature is intended to make communication between users easier, so please don't abuse it by posting duplicated, unsolicited @replies to lots of users. This is considered spam behavior!

In the automation section it says this

> Automated replies and mentions The reply and mention functions are intended to make communication between users easier, and automating these processes in order to reach many users is considered an abuse of the feature. If your application creates or facilitates automated reply messages or mentions to many users, the recipients must request or otherwise indicate an intent to be contacted in advance. For example, sending automated replies based on keyword searches is not permitted. Users must also have a clear and easy way to opt-out of receiving automated reply messages and mentions from your application. Note: media or brands using auto-response campaigns must request approval from Twitter and may be subject to additional rules. Please reach out to your account or partner manager for assistance. If you do not have a partner manager but are interested in managed client services, contact us.

Think we can provide a way to opt-out, but slightly worrying... Even if we get away with it, does it make it less policy relevant? Perhaps its evidence on the usefulness of twitter that is more experimentally clean than it would be if we didn't use mentions. Can possibly be extrapolated to ads

Ways around it:

- Doing tweets manually. 1000 tweets per day, say, between 5 people would be 10 minutes each or so.
- Setting up multiple accounts. This way we are less likely to be identified as mass tweeters
- Setting up multiple bots, and use tor to change the IP addresses.