Optimal Policies to Battle the Coronavirus "Infodemic" Among Social Media Users in Sub-Saharan Africa

Preanalysis plan

Molly Offer-Westort, Leah R. Rosenzweig, Susan Athey

July 23, 2020

Contents

ABSTRACT

Alongside the outbreak of Coronavirus, much of the world's population is also experiencing an "infodemic" – the spread of myths and hoax cures related to the virus. COVID-19 misinformation is spreading through online media outlets and social media platforms. While many false cures are largely harmless (e.g., drinking lemon water), others have devastating consequences, such as taking chloroquine. As a result, governments struggling to prepare healthcare systems and encourage citizens to comply with best practices also need to tackle misinformation. Building upon the experimental literature on combating fake news, we evaluate the effect of interventions designed to decrease sharing of false COVID-19 cures. Using Facebook advertisements to recruit social media users in Kenya and Nigeria, we deliver our interventions using a Facebook Messenger chatbot, allowing us to observe treatment effects in a realistic setting. Our aim is to find the context-aware intervention policy that will assign respondents the intervention that is most effective for them, based on their covariate profile. Using a contextual adaptive experimental design to sequentially assign treatment probabilities, we are able to learn the optimal contextual policy, and minimize assignment to ineffective or counter-productive interventions within the experiment. Analyzing heterogeneity in treatment effects allows us to learn whether different interventions are more effective for different people, improving our understanding of how to tackle harmful misinformation during an ongoing health crisis. Finally, we bring comparative data to a global problem for which the existing research has largely been limited to the U.S. and Europe. This pre-analysis plan describes the research design and outlines the key hypotheses that we will evaluate.

1. Motivation and Research Questions

Alongside the outbreak of Coronavirus, much of the world's population is also experiencing an "infodemic" – the spread of myths and hoax cures related to the virus. COVID-19 misinformation is spreading through online media outlets and social media platforms. These falsities range from incorrect information about government action to hoax cures. Unproven cures vary from mostly harmless, such as drinking lemon water, to those that can have devastating consequences if adopted, such as taking chloroquine or drinking bleach.[1] As a result, governments struggling to prepare health care systems and encourage citizens to comply with best practices are also struggling to tackle a pandemic of misinformation.

This project evaluates the effect of interventions designed to decrease sharing of false COVID-19 cures. Using Facebook advertisements to recruit social media users in Kenya

[1] For instance, dozens of people in Iran died from alcohol poisoning after ingesting methanol to stave off Coronavirus (Bloomberg News, Mar. 10, 2020).

and Nigeria, we deliver our interventions using a Facebook Messenger chatbot, allowing us to observe treatment effects in a realistic setting. We test interventions targeted at both the respondent level, such as general warnings, as well as headline-level treatments, such as flags and "false" tags (treatments are described in Table 1).

Using a contextual adaptive experimental design, we sequentially assign treatment probabilities to privilege assignment to the most effective interventions, and minimize assignment to ineffective or counter-productive interventions. Our aim is to learn an optimal contextual policy that will assign respondents the intervention that is most effective for them, conditional on their covariate profile. Exploring heterogeneity in treatment effects allows us to learn whether different interventions are more effective for different people, improving our understanding of how to tackle harmful misinformation during an ongoing health crisis.

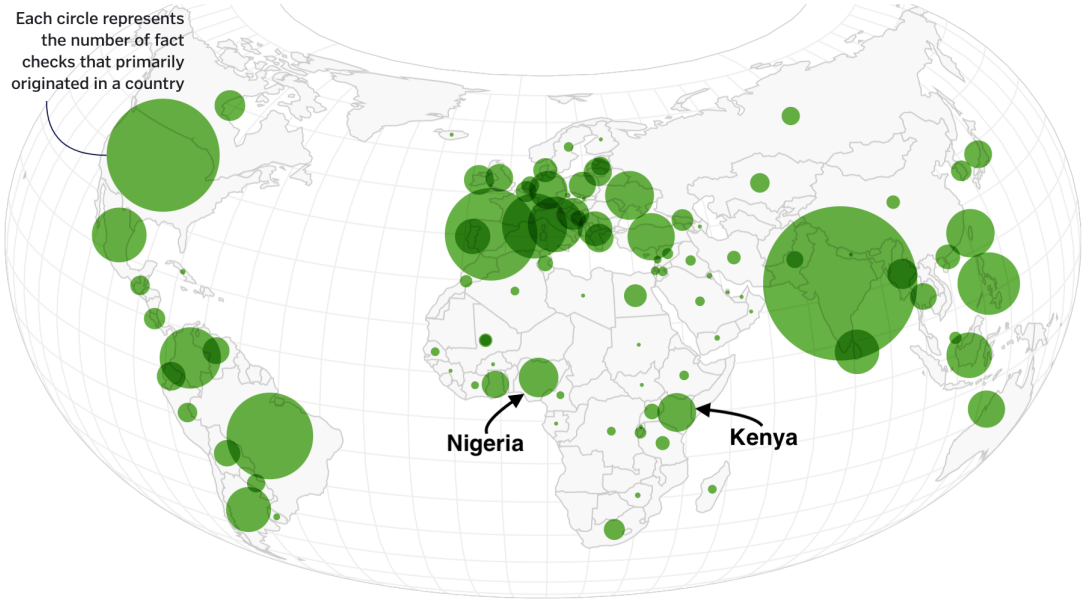
This work builds on the experimental literature on combating fake news in several important ways. First, we examine several prominent interventions that have proven successful in other studies and in other settings using an adaptive design to observe the best intervention policy. Second, we bring comparative data to a global problem. Despite the global nature of the "infodemic," much of the existing research has been focused on the Global North, particularly the United States (Pennycook et al., 2020; Bursztyn et al., 2020).[2] Finally, this pre-analysis plan describes the research design, outlines the key hypotheses that we will evaluate, and details our approach to analysis.

2. Case Selection

We examine these questions using a study focused on online social media users in two major English-language hubs of online communication in sub-Saharan Africa, Kenya and Nigeria. Collectively, Facebook estimates there are 30-35 million Facebook users who are 18 years and older from these two countries (as reported on Facebook's advertising platform). Misinformation and fake news are major problems in these countries. AfricaCheck.org, a third party verification site, has offices in both countries and has recently created pages devoted to COVID-19-related misinformation circulating online. From January to March, the number of English-language fact-checks increased by more than 900% worldwide (Brennen et al., 2020), demonstrating the prevalence of this kind of content and the availability of verified Coronavirus-related information. Figure 1 illustrates the volume of fact checks that appear in `poynter.org`'s global Coronavirus facts database, which demonstrates that Kenya and Nigeria are main factcheck sources on the continent. Thus, there is a large database of verified information from which we can draw stimuli for our experiment in these two countries.

[2] Two recent exceptions from sub-Saharan Africa include a field experiment in Zimbabwe using Whatsapp messages from a trusted NGO to counter COVID misinformation (Bowles et al., 2020) and a recent survey among traders in Lagos, Nigeria looking at the correlates of belief in COVID-related misinformation (Goldstein and Grossman, 2020).

Figure 1. Map illustrating the volume of fact-checks in poynter.org's global coronavirus facts database.

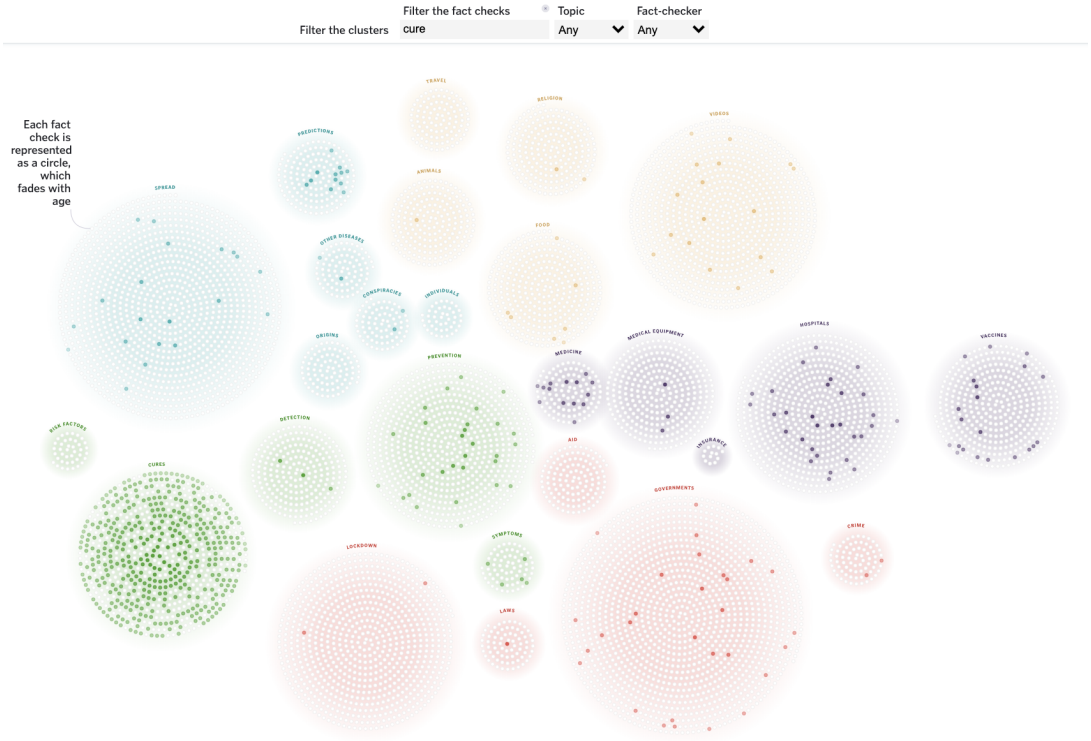


For this experiment, we focus on COVID-19 prevention and cure-related information because this comprises a large proportion of the overall Coronavirus-related information that has been fact-checked by experts (see Figure 2) and also serves as some of the most dangerous misinformation. Some hoax cures, when adopted, can be deadly. Moreover, even if not adopted when claims about the existence of a cure circulate widely they may deter people from taking preventative measures. We acknowledge that interventions will likely need to be specific to the particular type of misinformation being targeted, whether political, health-related, etc. The focus of this paper is on prevention and cure-related (mis)information that is currently relevant for the ongoing pandemic.

Should this be moved to a new 3.3.2 Stimuli section?

To collect stimuli we adopted several criteria to search for both false and true pieces of information related to Coronavirus prevention techniques and COVID-19 cures. First, we searched AFP, Poynter, and AfricaCheck website for any of this type of misinformation that had been checked by these organizations that appeared online in Kenya and Nigeria since the start of the pandemic in early March 2020. Second, we collected WHO myth-buster infographics that spoke to some of the misinformation we found. We also collected prevention messaging from the Nigeria Center for Disease Control, National Emergency Response Committee in Kenya, and the Ministry of Health in both countries, as these are the main government entities combating the spread of the disease in these countries and official sources of information. Our full set of stimuli for each country is presented in Appendix B.2.

Figure 2. Map illustrating the volume of COVID-19 cure-related fact-checks in poynter.org's global coronavirus database.



3. Experimental Setup

3.1. Sample recruitment

We will recruit respondents in Kenya and Nigeria using Facebook advertisements targeted to users 18 years and older living in these countries.[3] To achieve balance on gender within our sample we create separate ads targeting men and women in both countries. Our target sample size is 1,500 respondents in each country for our pilot. Size of the full scale study will be determined following piloting, in procedures described in Section 5.3.

Advertisements will appear within Facebook or Instagram, offering users with the opportunity to "Take a 15 minute academic survey on Messenger" in exchange for the equivalent of USD 0.50 in mobile phone airtime. When users click on the "Send Message" button on our advertisement, a Messenger conversation will open with our Facebook page, starting a conversation with a chatbot programmed to implement the survey. In contrast to sending users to an external survey platform such as Qualtrics, the benefit of the chatbot is that we keep users on the Facebook platform, with which they are likely more familiar, and maintain

Finalize advertisement text. LR: I edited - see appendix fig.

[3]Based on previous work it is clear that Facebook imputes location information for some of its users, which can be inaccurate. We will also ask a location screening question to ensure our respondents live in our countries of interest.

a realistic setting in which users might encounter online misinformation.[4] Respondents who complete the survey in the chatbot will receive compensation in the form of mobile phone airtime sent to their phone.

3.2. Covariates

Through the chatbot, we collect demographic and other information on respondents. We include the below covariates in analysis. The full list of covariates and question wording is in Appendix B.1.

Do we want to do policy analyses with all covariates? We can conduct this ex-post, we just can't update using the full model.

Covariate	**Coded as**
Gender	1 if male, 0 otherwise
Age	Indicators for population quartiles
Education	Indicators: No/informal schooling, any primary school, any secondary school, post-secondary qualifications, any university
Religion	Indicators: Muslim, Christian, other
Religiosity (freq. of attendance)	Indicators: less than a week, more than once a week but less than daily, daily
No. people in household	Indicators for population quartiles
Index of scientific views	Indicators for integers 0:2
Concern regarding COVID-19	Indicators: Very/somewhat/not worried
Perceived government efficacy on COVID-19	Indicators: Positive, neutral, negative

Discuss covariates. As is, this is 38,880 unique covariate combinations; is that feasible for implementation with zapier/google sheets/-bootstrapping? In practice, make sure we can run this process if we see a category we've never observed before.

3.3. Treatment

Drawing on the literature on experimental interventions to combat misinformation, we include several interventions designed to reduce the spread of misinformation online, which are targeted both at the headline level and respondent level. This list of treatments also draws on real-world interventions that companies and platforms have instituted to combat misinformation. Treatments are presented in Table 1.

[Table to be updated]

[4]The recruitment advertisement is shown in Figure 3 in Appendix A. [[TK: images of chatbot once linked to page]]

Shorthand Name	Treatment Level	Treatment
Facebook tips	Respondent	Facebook's "Tips to Spot False News"
AfricaCheck tips	Respondent	Africacheck.org's guide: "How to vet information during a pandemic"
Video training	Respondent	BBC Video training
Emotion suppression	Respondent	Prompt: "As you view and read the headlines, if you have any feelings, please try your best not to let those feelings show. Read all of the headlines carefully, but try to behave so that someone watching you would not know that you are feeling anything at all" (Gross, 1998).
Accuracy nudge	Respondent	Placebo headline: "Do you think this headline accurately describes an event that actually happened?" (Pennycook et al., 2020).
Deliberation nudge	Respondent	Placebo headline: "In a few words, please say *why* you would like to share or why you would not like to share this headline." [open text response]
Related articles	Headline	Facebook-style related stories: below story, show one other story which corrects a false news story
Factcheck	Headline	Fact checking flag from third party (e.g., Facebook, AFP, AfricaCheck, etc)
More information	Headline	Provides a link to "Get the facts about COVID-19" as per Twitter flags
Control	N/A	Control condition

Table 1. Description of interventions included in the experiment

Respondent-level treatments and headline-level treatments are implemented as separate factors, each of which has an empty baseline level that is the control. So respondents may be assigned the pure control condition, one of the respondent-level treatments but no headline-level treatment, one of the headline level treatments but no respondent-level treatment, or one of the respondent-level treatments *and* one of the headline-level treatments.

3.4. Outcomes and Response Function

We are interested in decreasing sharing of harmful false information about COVID-19 cures and treatments while not negatively impacting sharing of useful information about transmission and best practices from verified sources.

Primary Response Function We measure interest in sharing information through two questions:

- Would you like to share this post on your timeline?
- Would you like to send this post to a friend on Messenger?

Prior to treatment, we show respondents two articles randomly sourced from our misinformation stimuli and two articles randomly sourced from our true information stimuli, in random order, and for each stimuli we ask the above self-reported interest questions. Respondents are then asked a series of unrelated questions, and are then randomly assigned treatment according to the experimental design. If assigned one of the respondent-level treatments, they are administered the relevant treatment. They are then shown two additional misinformation stimuli and two additional true information stimuli, selected from the remaining stimuli that they were *not* shown pre-treatment. If the respondent is assigned a headline-level treatment, this treatment is applied only to the misinformation stimuli, as flags and fact-checking labels are not generally applied to true information from verified sources.[5] For each of the stimuli we again ask the same self-reported interest questions.

By using a pretest-posttest design [TK: citation] and an index of repeated measures (Broockman et al., 2017), we aim to improve the efficiency of our design.

We code response to the self-reported interest questions as 1 if the respondent affirms and 0 otherwise. Let M_i^1 be the sum of respondent i's pre-treatment responses to the *misinformation* stimuli and let T_i^1 be the sum of respondent i's pre-treatment responses to the *true* informational stimuli. M_i^2 and T_i^2 are the respective post-treatment responses. Then $M_i^1, T_i^1, M_i^2, T_i^2 \in 0,1,2,3,4$.

Our response function is then:

$$Y_i = -(M_i^2 - M_i^1) + 0.5 \times (T_i^2 - T_i^1)$$

Because of random assignment, we expect to see no systematic differences in interest in sharing either true or untrue stimuli across treatment conditions, conditional on covariates. For a given treatment condition, all else equal, if respondents share misinformation at lower rates post-treatment compared to control, this will result in a relatively higher response variable. If respondents share true information at lower rates post-treatment compared to control, this will result in a relatively lower response variable, but the relative impacts are only half as large as those for the misinformation stimuli.

This response function will be the variable which we optimize for in our adaptive algorithm described in Section 4.2, and in our policy learning described in Section 5.

LR:Made me wonder whether we want to, for power sake, use random effects model w/respondent and headline level REs? so rather than using DV that sums shares across all 4 headlines seen ? This is how psych misinfo studies usually analyze headline-level outcomes and deal with respondent and headline correlations.
MOW: Let's discuss, I'm not sure what this would look like for our adaptive agent model.

[5] The initial implementation of Twitter's labeling of coronavirus-related tweets with links to additional information was deemed to be overly broad, and was applied to some tweets that did not include misinformation. Twitter revised their labeling in late June of 2020. A company message was released here on June 26: https://twitter.com/TwitterSupport/status/1276661483561029632.

Secondary Outcomes Additionally, we measure several secondary outcomes, which we will also report.

Formalize how these secondary outcomes are analyzed, based on eventual below linear model.

These include willingness to read the article, as measured by the question: "If it were possible, would you click this headline to read the full story?" As above, we code responses as 1 if the respondent affirms and 0 otherwise. We separately calculate post-treatment minus pre-treatment differences for misinformation and true informational stimuli, and report estimates in the format [TK].

Additionally, we record behavioral measures. In order to obtain a behavioral measure of sharing, we collect the articles the respondent indicated they would like to share throughout the survey and at the end of the survey provide links to the *true* information. For these true stimuli, we offer respondents the opportunity to actually share this information as a Facebook post, which has been created on our project Facebook page. We are able to measure whether respondents click on a button which opens a pop-up screen to share the post on Facebook, however, we cannot measure directly whether they then actually follow through to the second step and post the article on their own timeline. Consequently, we report rates of clicking the initial share button in the format [TK]. Additionally, we report the *aggregate* number of times the associated post for each stimuli was shared.

We *could* actually create a separate post for each treatment combination, and measure sharing that way, but this seems like overkill for a secondary outcome in an already complicated project.

At this point we also debrief respondents, informing them about the headlines they were shown that are false. Instead of allowing respondents to share these headlines, we provide links to tips for spotting misinformation online and also offer them the opportunity to share these tips on their timeline or on messenger.

4. Hypotheses and Data Collection

Our data is described by treatments $W_i \in \mathcal{W}$[6]; response, $Y_i \in \mathbb{R}$; and covariates, $X_i \in \mathcal{X}$.

We assume the data is indexed by $i = 1, \ldots, N$ where indexing represents the order in which respondents entered the experiment; this allows us to use i to also represent relative chronological relationships in our sequential adaptive design.

We use potential outcome notation, where $Y_i(w)$ represents the potential outcome for respondent i under treatment w, and by experimental design, we have strong ignorability of potential outcomes to treatment conditional on observed covariates.

We would like to learn and evaluate an optimal contextual policy, under which we assign

[6] Our treatments are composed of two separate factors, but here we use W to represent combined treatment conditions. Where we wish to explicitly differentiate, we use W_i^R and W_i^H for respondent- and headline-level treatments respectively. As each factor includes a baseline level absent intervention, the cardinality $|\mathcal{W}| = |\mathcal{W}^H| \times |\mathcal{W}^R|$.

the most effective treatment conditional on covariates. Formally, a policy maps a set of covariates to a decision (Athey and Wager, 2017),

Update reference

$$\pi : \mathcal{X} \to \mathcal{W}. \tag{1}$$

In our setting, we will learn this policy, $\hat{\pi}$, and evaluate its value. The value of a policy is defined as,

$$V(\pi) = \mathrm{E}\left[Y(\pi(X_i))\right], \tag{2}$$

where the expectation is taken over the distribution of X.[7]

4.1. Hypotheses

Our hypotheses of interest relate the value of an estimated optimal contextual policy π_{opt} to fixed policies π_W, where under each fixed policy we would assign all respondents the relevant treatment w. The control policy is the fixed policy π_{w_C}

Our primary hypothesis is that we are able to estimate from the data an optimal contextual policy that improves over the control.

Hypothesis 1. *The best contextual policy that can be estimated from the data achieves higher value than the control treatment .*

$$H_0 : V(\pi_{opt}) = V(\pi_{w_C}) \qquad H_a : V(\pi_{opt}) > V(\pi_{w_C}) \tag{3}$$

This is the hypothesis that we aim to optimize for in our adaptive data collection.

We would also like to learn how much we gain by exploiting heterogeneity in the data. As a secondary hypothesis, we propose that the optimal policy that we are able to estimate from the data improves over the best fixed policy.

Hypothesis 2. *The best contextual policy that can be estimated from the data achieves higher value than the best fixed policy, i.e., the fixed policy with the highest associated value.*

$$H_0 : V(\pi_{opt}) = \arg\max_w V(\pi_w) \qquad H_a : V(\pi_{opt}) > \arg\max_w V(\pi_w) \tag{4}$$

[7] Here we will only consider deterministic policies, but for a random policy, the expectation will be taken over the joint distribution.

4.1.1. Provisional: additional pre-registered secondary hypotheses

LR: Do we want to specify that among the control group we will analyze the predictors of willingness to share fake news (maybe in this case more than true info, bc we want to predict ppl most susceptible to sharing falsities rather than just "always sharers")? – something we might hypothesize: belief in science/religiosity/COVID knowledge will be negatively correlated with sharing false cures... Age/trust in media will be positively correlated with sharing false cures. gender? [I dont have strong priors]. Unlike the US I think there will NOT be a strong correlation btw partisanship and willingness to share cures

MOW: I don't really have a position on this...happy to include if we think it is of interest.

4.2. Adaptive data collection

To collect data with the objective of learning an optimal policy, we use a *contextual bandit* algorithm, in which we sequentially update treatment assignment probabilities based on the observed history of treatments, response, and covariates. These types of algorithms navigate a tradeoff in *exploration* of the treatment space with *exploitation* of those treatments which we have observed to be effective based on historical data. This allows us to continue to learn about treatment effect heterogeneity while continuing to improve outcomes over time *within* the frame of the experiment.

We will use a version of linear Thompson sampling (Agrawal and Goyal, 2013). Under Thompson sampling (Thompson, 1933, 1935), treatment is assigned according to the Bayesian posterior probability that each treatment is best. In linear Thompson sampling, this is generalized to allow the outcome to be a linear function of covariates. We use a batched approach to updating, collecting data in batches and then updating treatment assignment model after each batch.

This algorithm requires us to se a number of design parameters; the procedures to select these parameters are discussed below in Section 5.3.

Our implementation is based on the balanced linear Thompson sampling method described in Dimakopoulou et al. (2017, 2019) :

Have James review notation and discuss

1. In the first batch, $b = 1$, we assign treatment uniformly at random.

The bootstrapping aspect of this is taken from the Cameroon PAP. Check wording and ensure that appropriate attribution is given.

2. For equally sized batches $b = 2, \ldots, B-1$:

    a) Fit a ridge regression model. Compute the minimum mean cross-validated error value of the penalization factor λ^{CV} using the entire observed history of data.[8]

Have James review model and discuss.

    b) Draw M random samples with replacement from the data, with samples indexed by $m = 1, \ldots, M$, so that data from sample m is represented by $D^{(m)} :=$

Option: Have separate penalty factors. Currently thinking one for just covariates, one for interactions of treatment x treatment and treatment x covariates jointly.

[8] For the agent we use a linear model, with treatment indicators, covariates, and treatment and covariates interacted:

$(X^{(m)}, W^{(m)}, Y^{(m)})$.[9]

c) Within each sample, for each possible context x, treatment w and bootstrap sample, estimate conditional means $\hat{\mu}_w^{(m)}(x)$ for each treatment using the fitted outcome model. The penalization value is fixed as λ_{CV} from step 1, without performing any additional cross-validation.

Can we simplify implementation here by using lasso instead of ridge, and collapsing our $\mathcal{X}$ to include only covariates that remain in the model?

d) For each context x and available treatment w, compute and store the following statistics representing the average value of each treatment, and the uncertainty associated with this statistic.

$$\begin{aligned}\hat{\mu}_w(x) &= \frac{1}{M}\sum_m \hat{\mu}_w^{(m)}(x) \\ \hat{\sigma}_w^2(x) &= \frac{1}{M(M-1)}\sum_m (\hat{\mu}_w^{(m)}(x) - \hat{\mu}_w(x))^2\end{aligned} \tag{6}$$

e) Approximate the probability that each treatment w is maximal for each possible context x. In order to do that, we draw from the following probability distribution a large number S times[10]

$$\theta_w(x) \sim \mathcal{N}(\hat{\mu}_w(x), \hat{\sigma}_w^2(x)) \qquad \text{for all treatments } w \tag{7}$$

and compute the fraction of times that treatment w was the largest for each s set of draws

$$q_b(x,w) = \frac{1}{S}\sum_s 1\left\{\theta_w^{(s)}(x) = \max\{\theta_1^{(s)}(x), \ldots, \theta_{|\mathcal{W}|}^{(s)}(x)\}\right\}. \tag{8}$$

These are the Thompson sampling probabilities associated with the pair (x,w).

$$\begin{aligned}\hat{\mu}_w(X_i) = &\sum_{w^R} 1\{W_i^R = w^R\}\hat{\beta}_{w^R} + \sum_{w^H} 1\{W_i^H = w^H\}\hat{\beta}_{w^H} + \\ &\sum_{w^R}\sum_{w^H} 1\{W_i^R = w^R\} \times 1\{W_i^H = \hat{w}^H\}\hat{\beta}_{w^R,H} + \\ &\sum_\ell X_{[\ell]i}\hat{\beta}_\ell + \\ &\sum_{w,\ell} 1\{W_i = w\}X_{[\ell]i}\hat{\beta}_{w,\ell}.\end{aligned} \tag{5}$$

The model is estimated using L_2 penalties for regularization, exclusive of the main treatment effects β_{w^R} and β_{w^R}. Observations are weighted according to stabilized inverse probabilities weights using known assignment probabilities, following Dimakopoulou et al. (2017). Stabilized inverse probability weights are discussed in Appendix C.1

[9] We set M = 100.

[10] We set S = 1,000.

f) Denote the control condition w_C, and assign a fixed probability $1/|\mathcal{W}|$ to the pure control condition, i.e., $\tilde{q}_b(x, w_C) = 1/|\mathcal{W}|$. For the remaining probabilities given each possible context x, update assignment probabilities so that they sum to 1, constraining the minimum assignment probability to a pre-determined probability floor, p

$$\breve{q}_b(x,w) = \max\left\{ \frac{q_b(x,w)}{\sum_{w \neq w_C} q_b(x,w)}, p \right\} \tag{9}$$

$$\tilde{q}_b(x,w) = \frac{\breve{q}_b(x,w)}{\sum_{w \neq w_C} \breve{q}_b(x,w)}. \tag{10}$$

g) Collect data for the batch: For every new respondent, collect data on their contexts x and use the probabilities $\tilde{q}_{b-1}(x,w)$ computed in the previous batch to assign treatments.

3. For the final batch, $b = \mathrm{B}$, collect data on-policy:

a) Estimate conditional means by fitting a random forest estimator on the entire data set collected through batch $B-1$, following the steps outlined in Appendix C.2, adjusting for adaptively collected data as described in Appendix C.3.

b) Fit a point-wise optimal policy by taking the maximum of predicted values for each possible context x

$$\hat{\pi}_x = \arg\max_w \hat{\mu}_w(x). \tag{11}$$

Store the policy.

c) Collect data for the batch: For every new respondent, collect data on their contexts, and assign treatment deterministically consistent with $\hat{\pi}_x$.

5. Analysis

Here, we can use a richer covariate set and e.g., information on stimuli. Specify.

To estimate the value of a policy, we take the average of doubly robust scores $\Gamma_{i,w}$, as in (12), following Robins et al. (1994)'s augmented inverse-propensity weighted scores.

$$\Gamma_{i,w} = \mu_w(X_i) + 1\{W_i = w\}\gamma_w(X_i)(Y_i - \mu_w(X_i)) \tag{12}$$
$$\mu_w(x) = \mathrm{E}\left[Y_i(w) | X_i = x\right]$$

We will estimate $\hat{\mu}_w(X_i)$ for each w using generalized random forests, following the approach is described in Appendix C.2. $\gamma_w(X_i)$ is a weight to account for unequal treatment assignment probabilities; we may use inverse probability weights calculated from the actual probabilities assigned under the experimental design; in practice, we use the stabilized versions of these weights , as described in Appendix C.1.

Could add note about bias in adaptively collected data, per e.g. Nie et al..

Our methods for analysis will differ depending on how the data is collected.

5.1. Policy learning and evaluation on randomly collected data

For randomly collected data, as in the pilot, we conduct policy learning and evaluation as below:

1. Collect data by assigning treatment uniformly at random.

2. Estimate nuisance components $\hat{\mu}_w(X_i)$ for each treatment separately, following the steps detailed in Appendix C.2; for $\hat{\gamma}_w(X_i)$, use assigned probabilities $1/|\mathcal{W}|$.

3. Compute doubly robust scores $\hat{\Gamma}_{i,w}$ substituting the estimated nuisance components into (12).

4. Fit a point-wise optimal contextual policy $\hat{\pi}_{opt}$ by taking the maximum of predicted values at each point
$$\hat{\pi}_{x_i} = \arg\max_w \hat{\mu}_w(x_i)$$

5. To evaluate the policies, take the average scores :
$$\hat{V}(\pi_w) := \frac{1}{N}\sum_i^N \hat{\Gamma}_{i,w}$$
$$\hat{V}(\hat{\pi}_{opt}) := \frac{1}{N}\sum_i^N \langle \hat{\pi}_{X_i}, \hat{\Gamma}_{i,\cdot} \rangle$$

Add a clarification on angle notation–or just change it, it's not obvious here.

6. To learn and evaluate the best fixed policy on a dataset, we cannot simply take the treatment condition with the highest estimated value, as this will give us positive bias

in expectation. To account for this, we use the approach described in Appendix C.4.

5.2. Policy learning and evaluation on adaptively collected data

For adaptively collected data, as in the simulations discussed in Section 5.3 and our eventual experiment, we conduct policy learning and evaluation as below:

1. Collect data under the adaptive algorithm described in Section 4.2.

2. For our nuisance components, due to the dependent nature of the data, we must ensure that our estimation is conducted using only historical data. Estimate nuisance components $\hat{\mu}_w(X_i)$ and $\hat{\gamma}_w(X_i)$ for data up to and including batch $B-1$ following the steps outlined in Appendix C.3.

3. Compute doubly robust scores $\hat{\Gamma}_{i,w}$ substituting the estimated nuisance components into (12).

4. We have already fitted and stored a point-wise optimal policy to conduct the on-policy evaluation in the final batch B of the adaptive experiment.

5. To evaluate the policies, we take the average scores over the relevant evaluation sets $\mathscr{I}$, where $\mathscr{I}_b$ represents the set of all observations within batch b. We note that evaluation of the optimal policy is simplified, due to the on-policy evaluation in the final batch B:

$$\hat{V}(\pi_w) := \frac{1}{\left|\bigcup_{b=1}^{B-1} \mathscr{I}_b\right|} \sum_{i \in \bigcup_{b=1}^{B-1} \mathscr{I}_b} \hat{\Gamma}_{i,w} \tag{13}$$

$$\hat{V}(\hat{\pi}_{opt}) := \frac{1}{|\mathscr{I}_B|} \sum_{i \in \mathscr{I}_B} Y_i \tag{14}$$

6. To learn and evaluate the best fixed policy on a dataset, we again take the relevant approach described in Appendix C.4.

To evaluate the hypotheses from Section 4.1, we estimate standard errors using the standard deviations of the relevant scores, and conduct frequentist hypothesis testing.

The data collected from this study will be used for eventual application of a contextual implementation of the evaluation weighting method proposed in Hadad et al. (2019), but those methods will not be discussed in this pre-registration.

Vitor and Ruohan: what should we say about confidence intervals paper?

5.3. Simulations and design parameters

Finalize how we're setting penalties

Note: *This section provides an overview of our approach to making data-driven design decisions. We will update this pre-analysis plan after collecting pilot data and running simulations, to document simulation results and our final design parameters, prior to implementing the eventual adaptive experiments.*

To carry out implementation, the above description requires setting of several design parameters, including total experiment size N, number of batches B, size of first batch $|\mathscr{I}_1|$, size of last batch $|\mathscr{I}_B|$, and probability floor p.

We set these parameters by learning from our pilot data of 1500 observations from each country. In the pilot data, treatment is assigned uniformly at random. We conduct the below simulations *separately* for each country, meaning that we may end up with meaningfully different designs in the two countries.

We then simulate data generating processes (DGPs) based on the pilot data, with varying heterogeneity. We create these DGPs by fitting a model to each dataset following (5), but instead of learning and applying the cross-validated penalty factor λ^{CV}, we generate models with varying complexity by over- and under-fitting to the data, imposing different penalty factors. In ridge regression, larger penalties will be associated with more parsimonious models, and less heterogeneity. Smaller penalties will be associated with more complex models, and consequently more heterogeneity. This approach allows us to generate heterogeneity that would plausibly exist in the true underlying populations.

Determine if our data generating model should be more complex. Discuss with James.

We refer to the heterogeneity "ratio" as the ratio of the value of the best contextual policy over the value of the best fixed policy. A ratio of two would indicate that the best contextual policy returns response that is in expectation twice as large as response under the best fixed policy. We can create a DGP with no heterogeneity by setting an arbitrarily large penalty factor, shrinking all treatment $\times$ covariate interactions to (effectively) zero.

James to review for notational consistency; refer to our implementation of the bandit agent.

Data generating processes

1. Define a vector of potential λ values, [TK, inclusive of zero].

2. Sample $M = 10,000$ observations with replacement from the empirical distribution of covariates in the pilot data; store this as $X^{(1)}, \ldots, X^{(M)}$.

3. Estimate heterogeneity ratios under each element the vector of penalty factors:

    a) Fit the model (5) to the pilot data under the relevant penalty factor to generate conditional means models $\mu_w(X)$ for each treatment w.

b) Calculate conditional means $\mu_w(X^{(m)})$ under the above fitted model conditional on covariates $X^{(1)}, \ldots, X^{(M)}$.

c) Estimate and store values for fixed policies for each w

$$\hat{V}(\pi_w) := \frac{1}{M} \sum_m \mu_w(x^{(m)})$$

d) Fit a point-wise optimal policy on the resampled data by taking the maximum conditional mean for each individual context $x^{(m)}$

$$\pi_{x^{(m)}} = \arg\max_w \mu_w(x^{(m)}). \tag{15}$$

e) Estimate and store value for the optimal policy:

$$\hat{V}(\hat{\pi}_{opt}) := \frac{1}{M} \sum_m \hat{\mu}_{\hat{\pi}_{x^{(m)}}}(x^{(m)})$$

f) Estimate the heterogeneity ratio as $\hat{V}(\hat{\pi}_{opt})/\hat{V}(\hat{\pi}_{w_{max}})$, where w_{max} is the true best arm under the relevant conditional means model over the empirical distribution of covariates.

4. Search over the vector of potential penalty factors to find:

a) The factor with an associated heterogeneity ratio that is closest in absolute distance to 1.05. This will allow us to learn about the performance of our algorithm in a case with a small amount of heterogeneity.

b) The largest penalty factor within one standard deviation of cross validated error from no penalization.

Does this make sense? Maybe plot hetereogeneity ratio and see what it looks like?

c) The two penalties factors which minimize the absolute distance to 1/3 and 2/3 of the distance between 1.05 and the above near-largest heterogeneity ratio.

Simulations This then gives us four conditional mean models. We then generate data from these models by:

1. Sampling covariates from the empirical distribution from the pilot data and assigning response as the conditional mean + a noise term, where the noise term is based on

the mean error between the fitted model and the pilot data, estimated separately for each treatment.

2. We run a series of simulated experiments using data from each of the DGPs, randomly applying parameters from Table 2 so that we have 500 iterations of experiments run under each unique combination of design parameters for each DGP.

Parameter choice Our objective in selecting design parameters is to optimize power for Hypothesis 1, while minimizing the size of the experiment and the number of batches. From the simulations we should be able to learn about power conditional on each combination of design parameters. Our decision rule is as follows:

1. Estimate power for Hypothesis 1 under each unique combination of design parameters for each DGP. Take the average power across DGPs, conditional on each unique set of design parameters.

2. If there is one or fewer combinations of design parameters with average power $\geq .8$, select the set of design parameters which optimizes Hypothesis 1. To break ties, select the set with smallest experiment size, or, if of equal size, select with smallest number of batches. If experiment size and batch size are equal, select randomly.

3. If there is more than one combination of design parameters with average power $\geq .8$, constrain choices to only those sets with average power $\geq .8$. Then constrain choices to only those sets with the smallest experiment size, and then to the smallest number of batches. Among the remaining sets, optimize for power of Hypothesis 1. To break ties, select randomly.

Table 2. Design parameters

Parameter	**Choice set**
First batch size	[100, 200, 400, 800]
Number of batches	[3, 4, 5]
Total experiment size	[2500, 3750, 5000]
Probability floor	$[0.01, 0.025, 0.5] \times \lvert\mathcal{W}\rvert$

Have James review design parameters and discuss.

References

Agrawal, S. and N. Goyal (2013). Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pp. 127–135.

Athey, S. and S. Wager (2017). Efficient policy learning. *arXiv preprint arXiv:1702.02896*.

Bowles, J., H. Larreguy, and S. Liu (2020). Countering misinformation via whatsapp: Evidence from the covid-19 pandemic in zimbabwe.

Brennen, J. S., F. M. Simon, P. N. Howard, and R. K. Nielsen (2020). Types, sources, and claims of covid-19 misinformation. *Reuters Institute*.

Broockman, D. E., J. L. Kalla, and J. S. Sekhon (2017). The design of field experiments with survey outcomes: A framework for selecting more efficient, robust, and ethical designs. *Political Analysis 25*(4), 435–464.

Bursztyn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott (2020). Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-44).

Dimakopoulou, M., S. Athey, and G. Imbens (2017). Estimation considerations in contextual bandits. *arXiv preprint arXiv:1711.07077*.

Dimakopoulou, M., Z. Zhou, S. Athey, and G. Imbens (2019). Balanced linear contextual bandits. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 33, pp. 3445–3453.

Goldstein, J. A. and S. Grossman (2020). Social media, partisanship, and covid-19 misinformation: Evidence from nigeria.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of general psychology 2*(3), 271–299.

Hadad, V., D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey (2019). Confidence intervals for policy evaluation in adaptive experiments. *arXiv preprint arXiv:1911.02768*.

Pennycook, G., J. McPhetres, Y. Zhang, J. G. Lu, and D. G. Rand (2020). Fighting covid-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological science*, 0956797620939054.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association 89*(427), 846–866.

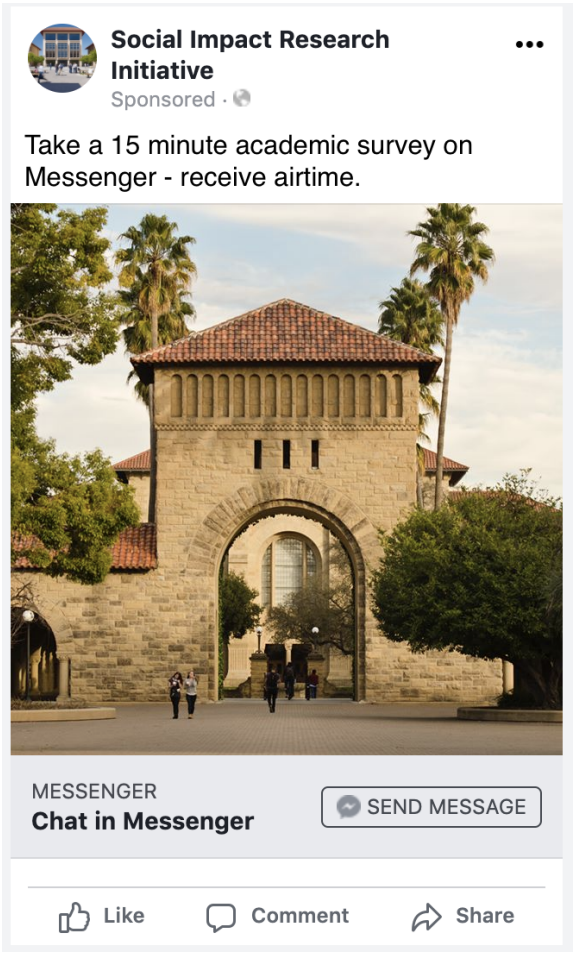
Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika 25*(3/4), 285–294.

Thompson, W. R. (1935). On the theory of apportionment. *American Journal of Mathematics 57*(2), 450–456.

Tibshirani, J., S. Athey, and S. Wager (2020). *grf: Generalized Random Forests*. R package version 1.2.0.

A. Recruitment

Figure 3. Advertisement as run in Facebook timeline.



LR: note I changed the recruitment text to not mention covid so we dont bias sample towards ppl that believe/interested in it or high info-seekers

B. Survey and data

B.1. Covariates

[TK: Full list of covariate questions and response options]

Optional: include link to full survey script? Would need to do some tidying

B.2. Stimuli

All of the stimuli used in the experiment are available at this link:
https://docs.google.com/spreadsheets/d/1ZEi8xU-TOZCZIQnDqq4VYjG5cWjIaWNyoKvPCjLL3fg/

edit?usp=sharing

B.3. Treatments

[TK]

Include mockups for headlines, full questions for respondent-level treatments, add deliberation stimuli and describe

C. Estimation Considerations

C.1. Stabilized inverse probability weighting

Inverse probability weighted estimation typically uses weights as follows,

$$\gamma_w^{IPW}(X_i) = \frac{1}{e_w(X_i)} \tag{16}$$
$$e_w(x) = \Pr[W_i = w | X_i = x].$$

Here, we could directly plug in the respective treatment assignment probabilities from the experimental design for the $e_w(X_i)$.

We use the stabilized version of this weights, normalizing weights to sum to one on the empirical data. This may improve RMSE of the estimator [TK:citation].

$$\gamma_w^{SIPW}(X_i) = \frac{1}{e_w(X_i)} \Bigg/ \sum_{j=1}^{N} \frac{1\{W_j = w\}}{e_w(X_i)} \tag{17}$$

For adaptively collected data, when we use stabilized weights in our algorithm, we use rolling weights that are updated with each batch.

Is this how we want to do this?

C.2. Random forest estimation

For policy learning and evaluation, we estimate conditional means using generalized random forests, as implemented by the `grf` package in `R` (Tibshirani et al., 2020).

For a given dataset, we estimate conditional means under each treatment condition w:

Have Erik review for terminology.

1. Fit a random forest estimator on the observations assigned w .

2. For observations assigned w, calculate $\hat{\mu}_w(X_i)$ using out-of-bag predictions.

3. For observation not assigned w, calculate $\hat{\mu}_w(X_i)$ using regression forest predictions from the model in step 1.

C.3. Adaptively weighted doubly-robust estimation

Add appropriate reference to LFO paper.

For adaptively collected data, we use doubly robust scores as in (12), but due to the dependent nature of the data, to avoid bias, we must ensure that we use only historical data in our estimates. This means that in each batch we estimate the nuisance components only using data up to and including the current batch.

To estimate conditional means, we follow the steps above in C.2, with minor adjustments. For each batch b in $b = 1, \ldots, B-1$ and for each treatment w:

1. Fit a random forest estimator on the observations assigned w in batches up to and including batch b.

2. For observations assigned w in batch b, calculate $\hat{\mu}_w(X_i)$ using out-of-bag predictions.

3. For observation not assigned w in batch b, calculate $\hat{\mu}_w(X_i)$ using regression forest predictions from the model in step 1.

We use the rolling version of the stabilized inverse probability weights from (17), substituting the current maximum index value for N. Doubly robust scores are then formed from the relevant component parts.

C.4. Random best fixed policies

We are interested in learning and evaluating the best fixed policy. However, if we learn which fixed policy is best by taking the fixed policy with the highest mean, we get a biased estimate of the best fixed policy. To see this, consider:

$$\mathrm{E}\left[\max(X_1, \ldots, X_N)\right] \geq \max(\mathrm{E}\left[X_1\right], \ldots, \mathrm{E}\left[X_N\right]).$$

To address this concern, we consider instead a *random* best fixed policy.

1. For each observation $i > 1$ in the experiment, we calculate the value of fixed policies

as the average of scores up to time $i-1$.

$$\hat{V}_{i-1}(\pi_w) := \frac{1}{i-1}\sum_{j=1}^{i-1}\hat{\Gamma}_{j,w} \quad \text{for fixed policies } w$$

2. The "best" fixed policy in period i is the treatment with the highest estimate:

$$w_i^* = \arg\max_w \hat{V}_{i-1}(\pi_w)$$

3. The score for the random best fixed policy in time i is then the score in that period for the selected arm, $\hat{\Gamma}_{i,w^*}$

4. To evaluate the policies, we again take the average scores. The evaluation set $\mathscr{I}^*$ will be the entire data set for data collected under the procedures for the random agent as described above in Section 5.1, and up through batch $B-1$ for data collected under the procedures for the adaptive agent–excluding the first observation.

$$\hat{V}(\hat{\pi}_{w^*}) := \frac{1}{\left|\mathscr{I}^*\right|}\sum_{i\in\mathscr{I}^*}\hat{\Gamma}_{i,w_i^*}$$